The *Proceedings*

OF

THE INSTITUTION OF ELECTRICAL ENGINEERS

FOUNDED 1871 : INCORPORATED BY ROYAL CHARTER 1921

PART B

RADIO AND ELECTRONIC ENGINEERING
(INCLUDING COMMUNICATION ENGINEERING)

The Difficult and The Easy

The making of *some* Massicore transformers calls for the solution of new problems, and Savage of Devizes are continually pioneering techniques. Others are relatively simple to construct, for many industries can still use the well-established types of transformers.

But complicated or simple, the essential ingredient remains the same—conscientious craftsmanship. Savage Massicore transformers are built to last a lifetime.

Equally important is the individual attention given to all enquiries and orders regardless of size; and we make a point of keeping our delivery promises.
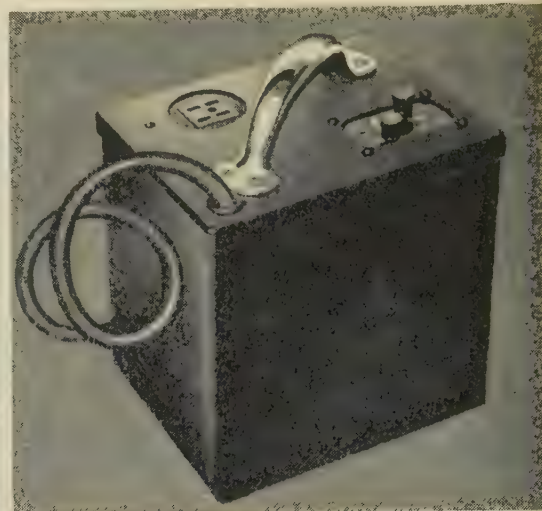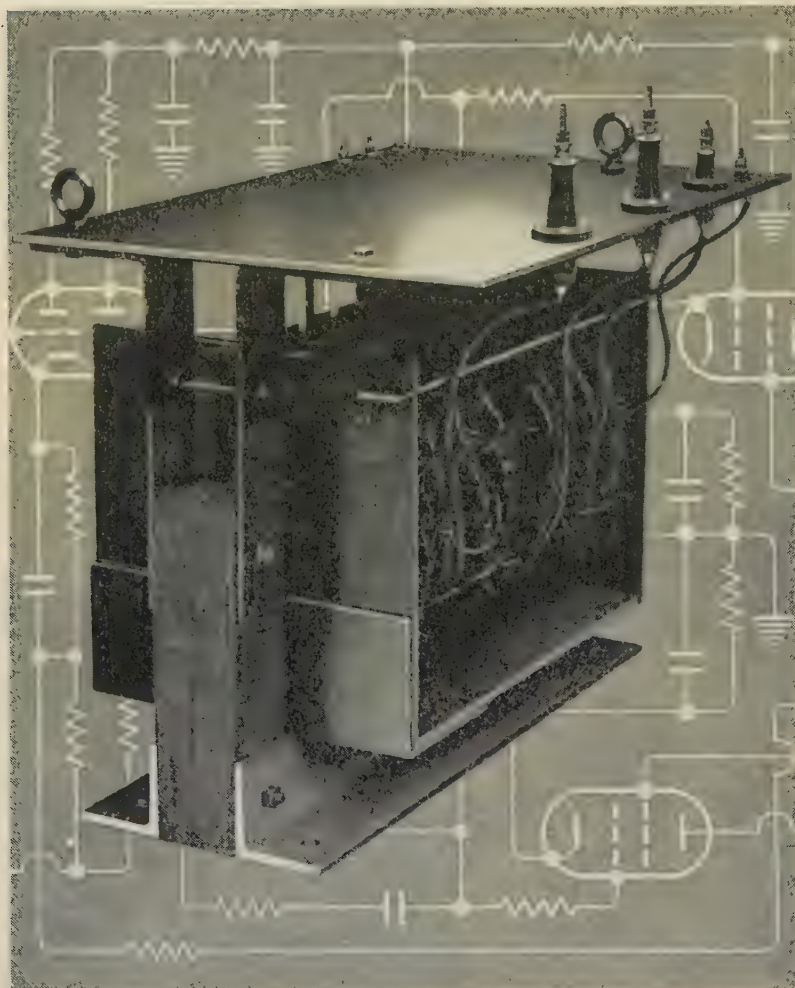
The special instrument on the left, far from orthodox in many ways, is just over two feet long and gives 250 watts at 2,500, 5,000 or 10,000 volts at frequencies between 400 c/s to 10,000 c/s.

And above, in contrast, is a simple isolating transformer for use where portable power supply is required, such as for hand tools, sub-standard film projectors, lighting etc., or for stepping mains voltage up or down.

Your requirements may call for instruments very different from these examples. Please take advantage of our experience, knowledge and constructional skill in the production of all types of transformers.

Corner for Contented Customers

"Thank you for your execution of our order for transformers and chokes to specification. I feel sure that the generous proportions v. rating of these components will ensure a long life for them.
It is refreshing to deal with a company which keeps to its delivery dates and I am pleased to express my appreciation."

N. J. C. BIRMINGHAM

b

The market-place of a Roman town A.D.350.

THEY SET A STANDARD

IN the steps of the Roman legions that conquered Britain came the merchants and settlers. Secure under Roman law and administration, commerce and agriculture flourished, and Ancient Britain experienced a period of peace and prosperity it was not to know again until long after the Dark Ages.

In government, as in many other fields, the Romans set a standard which few have equalled since.

In cable making too, standards are of vital importance. For over 100 years members of the Cable Makers Association have been concerned in all major advances in cable making.

Together they spend over one million pounds a year on research and development. The knowledge gained is available to all members.

This co-operation has contributed largely to the world-wide prestige that C.M.A. cables enjoy, and it has put Britain at the head of the world cable exporters. Technical information and advice is freely available from any C.M.A. member.

MEMBERS OF THE C.M.A.

British Insulated Callender's Cables Ltd. · Connollys (Blackley) Ltd.
Enfield Cables Ltd. · W. T. Glover & Co. Ltd. · Greengate &
Irwell Rubber Co. Ltd. · The Hackbridge Cable Co. Ltd.*
W. T. Henley's Telegraph Works Co. Ltd. · Johnson & Phillips Ltd.
The Liverpool Electric Cable Co. Ltd. · Metropolitan Electric Cable
& Construction Co. Ltd. · Pirelli-General Cable Works Ltd.
(The General Electric Co. Ltd.) · St. Helens Cable & Rubber Co. Ltd.
Siemens Edison Swan Ltd. · Standard Telephones & Cables Ltd.
The Telegraph Construction & Maintenance Co. Ltd.
C.M.A. Trade Marks for Mains Cables only

Insist on a cable with the C.M.A. label

The Roman Warrior and the letters 'C.M.A.' are British Registered Certification Trade Marks.

CABLE MAKERS ASSOCIATION

CABLE MAKERS ASSOCIATION, 52-54 HIGH HOLBORN, LONDON, WC1 TELEPHONE HOLBORN 7633

CMA 21

Publications of
THE INSTITUTION OF ELECTRICAL ENGINEERS

Proceedings of The Institution

PART A (Power Engineering)—Alternate Months

PART B (Radio and Electronic Engineering—including Communication Engineering)—Alternate Months

PART C (Institution Monographs)—In collected form twice a year

Special Issues

VOL. 94 (1947) PART IIA (Automatic Regulators and Servomechanisms Convention)

VOL. 94 (1947) PART IIIA (Radiocommunication Convention)

VOL. 97 (1950) PART IA (Electric Railway Traction Convention)

VOL. 99 (1952) PART IIIA (Television Convention)

VOL. 100 (1953) PART IIA (Symposium of Papers on Insulating Materials)

Heaviside Centenary Volume (1950)

Thermionic Valves: the First Fifty Years (1955)

VOL. 103 (1956) PART B Supplements 1–3 (Convention on Digital-Computer Techniques)

VOL. 103 (1956) PART A Supplement 1 (Convention on Electrical Equipment for Aircraft)

VOL. 104 (1957) PART B Supplement 4 (Symposium on the Transatlantic Telephone Cable)

VOL. 104 (1957) PART B Supplements 5–7 (Convention on Ferrites)

VOL. 105 (1958) PART B Supplement 8 (Symposium on Long Distance Propagation above 30 Mc/s)

VOL. 105 (1958) PART B Supplement 9 (Convention on Radio Aids to Aeronautical and Marine Navigation)

Science Abstracts

Section A: Physics—Monthly

Section B: Electrical Engineering—Monthly

Cumulative Index

Prices of the above publications on application to the Secretary of The Institution, Savoy Place, W.C.2
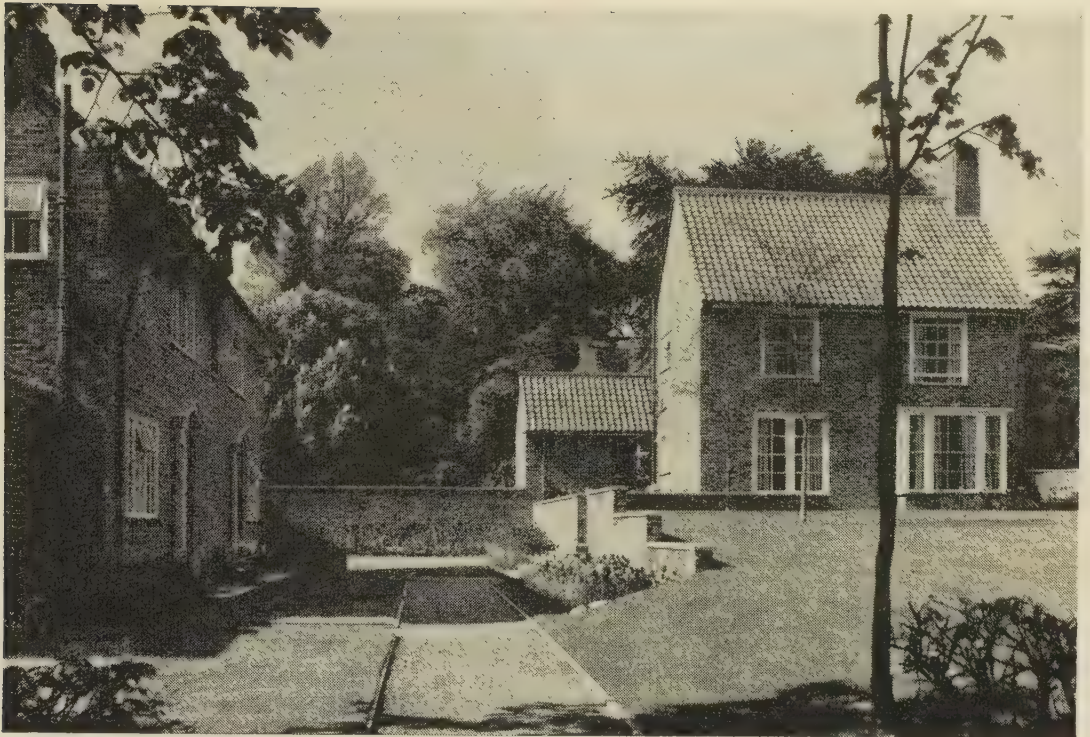
the homes fund

'The Chesters', at New Malden, Surrey,

is a residential estate for members of The Institution or their

dependants whose needs have come to the notice of the

Court of Governors of the Benevolent Fund.

Funds are still needed to complete the original scheme of 26

residences.

CONTRIBUTIONS HOWEVER SMALL ARE WELCOMED AND MAY BE SENT TO THE HON. SECRETARY OF THE INCORPORATED BENEVOLENT FUND, THE INSTITUTION OF ELECTRICAL ENGINEERS, SAVOY PLACE, LONDON, W.C.2, OR HANDED TO ONE OF THE LOCAL HON. TREASURERS OF THE FUND.

INDEX OF ADVERTISERS

DISCUSSION ON
'FREQUENCY-MODULATED V.H.F. TRANSMITTER TECHNIQUE'*

SOUTHERN CENTRE AT SOUTHAMPTON, 6TH NOVEMBER, 1957

Ir. G. Shearing: Until recently, little use has been made of quency modulation for communication. It seems para-tical that frequency modulation, which produces many side-ds compared with the two of amplitude modulation, can give her fidelity and signal/noise ratio in reception. The earlier erimental work in Europe and America has had excellent tion in the v.h.f. system described.

s there any evidence of multi-path indirect interference ween the service areas and is such interference likely to rease with the rapidly expanding European f.m. services?

Legarding the new method of transmitter control, does a shut-wn occur in the event of failure of any one of the supplies ch may not cause an immediate overload, e.g. air cooling? e the dual basic major elements interchanged, as required, by uote control? Has the air-cooling system experienced trouble m the filters or the valve insulation in the contaminated air ch may occur in some districts?

Ir. E. Wolfendale: When describing the centre-frequency trol circuits the authors express a preference for a binary quency-divider as opposed to the tuned-circuit divider actually d. The binary divider could not be incorporated in the asmitter, owing to the large number of valves required, but ald have been acceptable had it been possible to perform the e operation with transistors. A frequency divider operating n 7·5 Mc/s is possible using high-frequency alloy-junction isistors of the type OC44 or equivalent, and the method of ieving this has been described.† If the method described in paper is applied to the OC44 type transistor, then scaling uits with resolutions up to 0·1 microsec may be achieved. should be interested to learn the authors' reasons for pre-ing the binary divider.

Iessrs. A. C. Beck, F. T. Norbury and **J. L. Storr-Best** (*in y*): In reply to Mr. Shearing, interference between trans-ters in separate service areas has, in fact, been reported. It orth noting, however, that, in anticipation of the possible cts of this type of interference, the f.m. stations were so nned that at least a 20 dB ratio between the wanted and

unwanted carrier signals between co-channel transmitters would be secured in the corresponding service areas. This is without taking into account the improvement to be gained by the use of a directional aerial. Furthermore, it is not expected that the expansion of European f.m. services will give rise to carrier ratios poorer than this figure. Experiments have shown that a receiver with a discriminator of normal bandwidth and with an a.m. suppression ratio of 30 dB or better will be rarely affected by this order of interference. Although this a.m. suppression is con-sidered appropriate for a fairly good domestic f.m. receiver, inevitably some fall short of this target, and under such special conditions of interference, it must be expected that these will show some deterioration in reproduction.

The method of transmitter control was developed as a simple on/off system with a view to the special requirements of un-attended operation. Thus the failure of any section likely to affect the safety or serviceability of a transmitter is made to cause that transmitter to shut down. Air cooling is included in this category. The remote-control system provides for the complete change-over from one modulated drive to its standby; the transmitters are, however, normally running in parallel, so that no facilities are necessary for the interchange of transmitters in service.

On the question of contaminated air, this varies greatly in severity between different sites. In industrial areas the condi-tions are particularly adverse and frequent maintenance is necessary to keep the air filters clean. This problem is receiving study at the present time with a view to reducing maintenance and increasing filter life.

In reply to Mr. Wolfendale, the particular advantage of a binary frequency-divider in the present application would be the elimination of frequency-sensitive circuits. A true counting device would obviate the necessity for a second crystal standard in the frequency-control system, since the same device could be fed directly from the master oscillator irrespective of the par-ticular channel frequency. The problems of accurate alignment and reliable starting usually associated with tuned regenerative dividers would also be avoided. These advantages would, however, have to be weighed against the fact that a binary divider requires more semi-conductors, whether these be valves or transistors, than the tuned regenerative type.

Be k, A. C., Norbury, F. T., and Storr-Best, J. L., Paper No. 2267 R, November, (o 104 B, p. 225).
.......PLIN, G. B. B. and Owens, A. R.: 'A Junction-Transistor Scaling Circuit with .....: Resolution', *Proceedings I.E.E.*, Paper No. 2076 R, March, 1956 (103 B,

535—1

The Institution of Electrical Enginee
Paper No. 26(
July 19!

THE FORTY-EIGHTH KELVIN LECTURE

'INFRA-RED RADIATION'

By G. B. B. M. SUTHERLAND, M.A., Sc.D., F.R.S.

HISTORICAL

Very early in 1800, Sir William Herschel[1] began 'an investigation of the powers of the prismatic colours to heat and illuminate objects'. On the 8th March he wrote up his results, and the paper was communicated to the Royal Society on the 27th. It began as follows:

> It is sometimes of great use in natural philosophy to doubt of things that are commonly taken for granted; especially as the means of resolving any doubt, when once it is entertained, are often within our reach. We may therefore say, that any experiment which leads us to investigate the truth of what was before admitted upon trust, may become of great utility to natural knowledge. Thus, for instance, when we see the effect of the condensation of the sun's rays in the focus of a burning lens, it seems to be natural to suppose, that every one of the united rays contributes its proportional share to the intensity of the heat which is produced; and we should probably think it highly absurd, if it were asserted that many of them had but little concern in the combustion or vitrification, which follows, when an object is put into that focus. It will therefore not be amiss to mention what gave rise to a surmise, that the power of heating and illuminating objects, might not be equally distributed among the variously coloured rays.
>
> In a variety of experiments I have occasionally made, relating to the method of viewing the sun, with large telescopes, to the best advantage, I used various combinations of differently-coloured darkening glasses. What appeared remarkable was, that when I used some of them, I felt a sensation of heat, though I had but little light; while others gave me much light, with scarce any sensation of heat. Now, as in these different combinations the sun's image was also differently coloured, it occurred to me, that the prismatic rays might have the power of heating bodies very unequally distributed among them; and, as I judged it right in this respect to entertain a doubt, it appeared equally proper to admit the same with regard to light. If certain colours should be more apt to occasion heat, others might, on the contrary, be more fit for vision, by possessing a superior illuminating power. At all events, it would be proper to recur to experiments for a decision.

Using the sun as source and a glass prism to produce the spectrum, Herschel placed a thermometer successively in the various parts of the spectrum, finding a rise of about 7° in the red, 3° in the green and 2° in the violet. This proved that radiant heat was 'refrangible', but Herschel also noticed that the maximum of visibility in the spectrum (which lies in the green) did not coincide with the maximum of the radiant heating effect, which he had proved to be in the red. Moreover, he suspected that he might not have reached the maximum of the heating effect. Accordingly, he began a second series of experiments using the apparatus shown in Fig. 1.

He placed his thermometer on this occasion just beyond the red end of the spectrum and found a rise of 9° when the thermometer was placed 'half an inch out of the visible rays of the sun'. He concluded as follows:

> We have traced these calorific rays throughout the whole extent of the prismatic spectrum; and found their power increasing, while their refrangibility was lessened, as far as to the confines of red-coloured light. But their diminishing refrangibility, and increasing power, did not stop here; for we have pursued them a considerable way beyond the prismatic spectrum, into an invisible state, still exerting their increasing energy, with a decrease of refrangibility up

Dr. Sutherland is at the National Physical Laboratory.

to the maximum of their power; and have also traced them to tha state where, though still less refracted, their energy, on account, w may suppose, of their now failing density, decreased pretty fas after which, the invisible thermometrical spectrum, if I may so call i soon vanished.

These experiments, establishing the existence of invisible hea radiation, were completed and written up by the 17th Marcl 1800, and communicated to the Royal Society on the 24th Apri He had completed and written up a third series of experiment by the 26th April, the results of which were presented to th Royal Society on the 15th May. The speed with which h worked and was able to get his results before his fellow scientist should be better known as an object lesson to many of us.

In the third set of experiments he proved the properties o reflection and refraction for invisible heat rays and also measure their absorption in a wide variety of materials. These wer mostly glasses presumably, since it was his original interest i using coloured glasses in telescopes which had led him into thi field; but he also examined liquids—including well-water, se water (from the head of the pier at Ramsgate at high tide), win gin and brandy. These results are summarized in Table 1.

Table 1

HERSCHEL'S RESULTS ON THE ABSORPTION OF NEAR INFRA-REI
RADIATION BY VARIOUS LIQUIDS

Empty tube	and 2 glasses stop 542 rays of heat, and 204 of light
Spring water	,, ,, ,, ,, 558 ,, ,, ,, ,, 211 ,, ,,
Sea water	,, ,, ,, ,, 682 ,, ,, ,, ,, 288 ,, ,,
Spirit of wine	,, ,, ,, ,, 612 ,, ,, ,, ,, 224 ,, ,,
Gin	,, ,, ,, ,, 739 ,, ,, ,, ,, 626 ,, ,,
Brandy	,, ,, ,, ,, 794 ,, ,, ,, ,, 996 ,, ,,

It must be remembered that in 1800 the concept of heat wa still rather indefinite. Herschel took great care to call the subjec of his researches 'the rays that occasion heat', so that he coul not be 'misunderstood as meaning that the rays themselves wer heat'. Nor did he attempt to show in what manner they produce heat, but he said he was inclined to believe that the heat-makin rays 'occasioned vibrations in the parts of bodies'—a remarkabl example of scientific intuition. If Herschel had had at hi disposal a detector of infra-red radiation superior to a mercury in-glass thermometer, there is little doubt that he would hav gone on to prove the correctness of his speculation and show that infra-red rays do excite characteristic vibrational frequencie in molecules. The thermometer is such a crude detector tha Herschel was unable to penetrate into the most interestin region of the infra-red spectrum.

To make this point clearer, one has only to consider th distribution of the energy emitted by a hot 'black body' (i.e perfect emitter) over the electromagnetic spectrum. This illustrated in Fig. 2. For instance, when the temperature of th hot body is $2\,000°$K, most of the energy is concentrated ju beyond the red end of the visible spectrum, i.e. between $0·8$ and $3·5\mu$. Herschel's work was confined to that region, fir because of the insensitivity of the thermometer, and second because a glass prism does not transmit infra-red radiatio
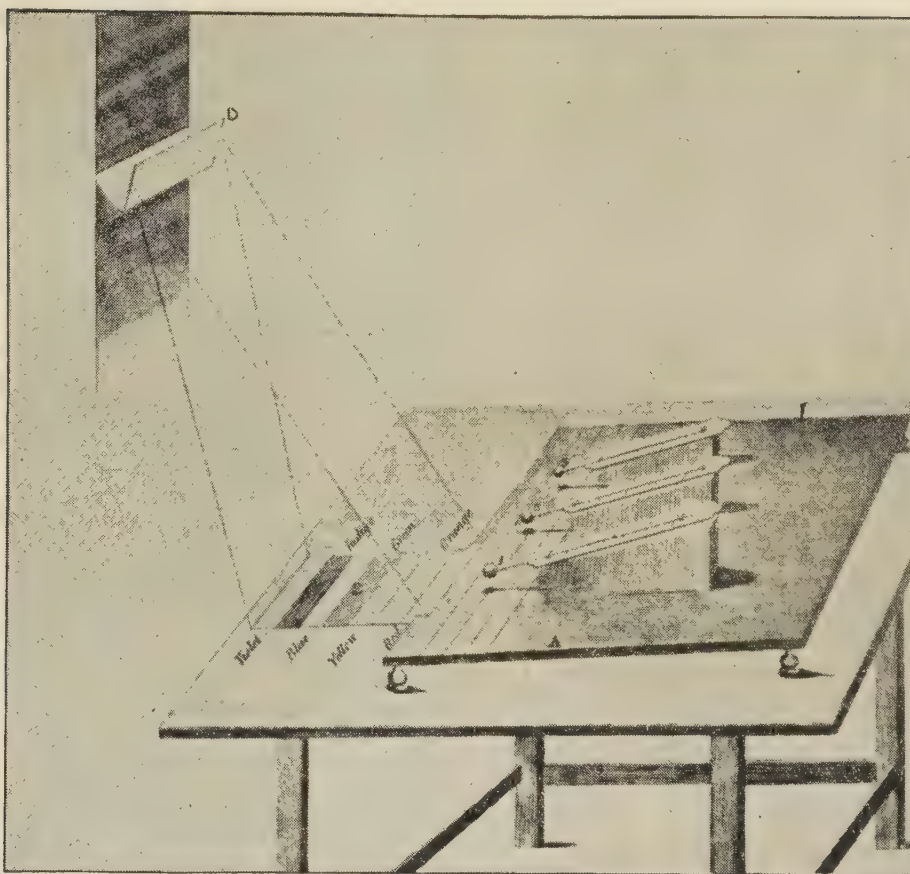
Fig. 1.—Apparatus used by Sir William Herschel in his pioneer work on infra-red radiation.

Solar radiation was dispersed by the prism D, and the distribution of heat radiation in this spectrum was investigated by observing the temperature rise in thermometers placed in and beyond the visible part of the spectrum.

ond about 3μ. Unfortunately, it is just between 3μ and out 30μ that the majority of the fundamental characteristic rations of molecules occur. Moreover, in order to study m, it is necessary to disperse the spectrum quite widely and xplore it in very small steps, thus reducing the energy falling the detector to a minute fraction of that which produces a ctable rise in a mercury-in-glass thermometer. It should be remarked that, as the temperature of the hot body drops, total energy emitted decreases very sharply (according to a aw) and the position of the maximum energy moves gradually onger wavelengths (Fig. 2). Progress in the infra-red region he spectrum has therefore been strongly dependent on the blopment of sensitive detectors.

DEVELOPMENT OF DETECTORS

here are an innumerable number of ways of detecting infra- radiation (including the subjective one with which our ation of heat is associated), but those which have been nd of scientific interest and value may be divided into twelve es. They are listed in Table 2, where they have been arranged reer of penetration into the infra-red.

will be noticed that the first six detectors are classed as todetectors, while the second six are called thermodetectors. distinction is a fundamental one, since it arises from their le of action and naturally affects their characteristic pro- ies. In photodetectors, the incident photon (or quantum) a-red radiation interacts with an electron in the detector mediately produces an observable effect; e.g. it changes

Table 2

Type of detector	Useful range
Photodetectors { (a) Photochemical	
(b) Phosphorescent	To about $1 \cdot 3\mu$
(c) Photoemissive	
(d) Photovoltaic	To about 5μ
(e) Photoelectromagnetic	To about 7μ
(f) Photoconductive	To about 8μ
Thermodetectors { (g) Evaporographic	To about 10μ
(h) Thermoelectric	All wavelengths
(i) Bolometric	All wavelengths
(j) Pneumatic	All wavelengths
(k) Radiometric	All wavelengths
(l) Thermometric	All wavelengths

the energy state of the electron so that it can take part in con-duction (photoconductor) or may even be emitted from the surface of the detector (photo-emitter). In thermodetectors the action is not so direct. In this case, infra-red radiation is absorbed by the detector, thus raising its temperature. This rise of temperature may affect the resistance of the detector strip (bolometer) or the pressure of a gas in thermal contact with the detector surface (pneumatic). Since the thermodetector depends on a heating effect and it is not difficult to 'blacken' any detector surface so that it absorbs all wavelengths, the thermo-detector is, in principle, sensitive to all wavelengths, although in some cases the effect may become undetectable at very long wavelengths (evaporographic limit, about 10μ). The photo-detector necessarily has a restricted range of operation. An incident photon of infra-red radiation has a specific amount of energy which is inversely proportional to the wavelength. If
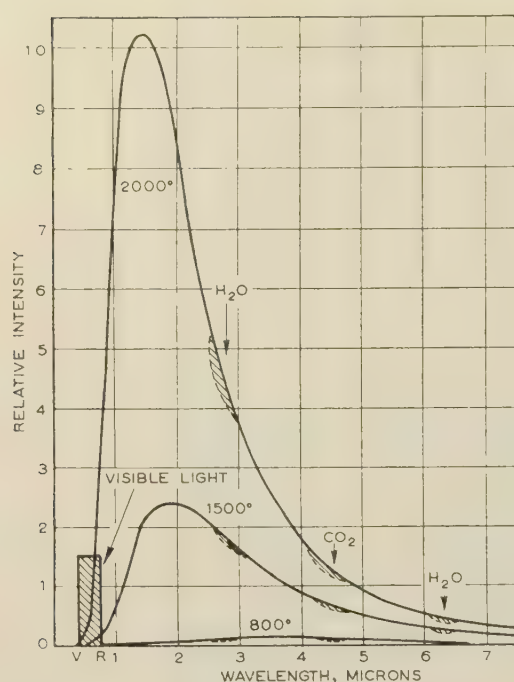
Fig. 2.—Distribution of the energy in the spectrum of a black body at various temperatures.

The visible part of the spectrum is the shaded rectangle between V and R. Atmospheric absorption due to CO_2 and H_2O is seen near $2\cdot8\,\mu$, $4\cdot3\,\mu$ and $6\cdot3\,\mu$. The quantity λ in the upper table is the wavelength of the maximum of the black-body curve.

$$\lambda T = C = 2\,940$$

	Room temp.	Red hot	White hot	Tungsten	Sun
T, deg K	300	800	1 800	3 200	5 880
λ, microns ..	9·7	3·68	1·63	0·92	0·5

For 2 000° K black body

λ (microns)	Relative intensity
1·44	1·0
2·00	0·8
5·00	0·09
10·00	0·009
50·00	0·000 02
100·00	0·000 001
200·00	0·000 000 08

the energy of the photon is too low to activate the electron, the photodetector cannot operate. This is why all photodetectors have a long wavelength limit beyond which they are completely insensitive. It should be added that the response of the photodetector usually varies for the different wavelengths to which it is sensitive (cf. photographic plate), whereas the thermodetector has a uniform response. One further general distinction is that photodetectors usually have a much faster response time than thermodetectors, since the latter require time for the energy in the absorbed quanta to produce a detectable change in temperature and for this heat to be removed so that the detector is ready to respond to a fresh impulse.

Thermodetectors

The thermometric and radiometric methods of detecting infrared radiation will not be discussed, since these are now only of historical interest. The first effective thermodetector was the thermocouple. This is still one of the most widely used detectors, for it has been improved enormously since its introduction by Melloni[2] in 1833 as an application of the thermo-electric effect,

discovered by Seebeck in 1822. The early thermocouples and thermopiles (made from a series of thermocouples) were neither very sensitive nor very fast. The modern thermocouple, made from carefully chosen semi-conductor alloys, can detect an incident power of less than 10^{-10} watt/mm², and has a time constant of less than $0\cdot01$ sec. It is of some interest that the best one is due to the late Dr. E. Schwarz, who came to this country as a refugee and was given the opportunity to develop this detector by Adam Hilger.

The thermocouple maintained its supremacy for fifty years, i.e. until 1880, when Langley[3] introduced the bolometer, in which the change in resistance of a fine wire is measured when its temperature is raised by the absorption of infra-red radiation. This discovery has been commemorated in the quaint limerick which goes:

> Mr. Langley invented the bolometer,
> Which is really a kind of thermometer
> That will measure the heat
> From a polar bear's seat
> At a distance of half a kilometre.

Although Langley's bolometer was superior to existing thermocouples, the latter were gradually improved, so that by 1905 when Coblentz[4] was doing his classical pioneer investigations on infra-red spectra, there was little to choose between them. The position is much the same to-day. In general, the thermocouple is rather more sensitive when one is working near the limits of detection, while the bolometer has a faster time of response (especially the thermistor bolometer). There is, however, one rather esoteric form of bolometer which is both faster and more sensitive than the best thermocouple. This is the superconducting bolometer developed by Andrews, Melton and de Sorbo[5] during the Second World War. It depends for its action on the transition to a near superconducting state of columbium nitride near $14\cdot3°$ K. It requires liquid hydrogen for its operation and is so elaborate and expensive that it has not yet become a practicable proposition, even in the research laboratory.

A more practical thermodetector, also developed during the Second World War, is the pneumatic detector, or Golay cell, named after its inventor Marcel Golay.[6] In this case, the infra-red radiation causes expansion in a small volume of gas, which expansion can be measured by cleverly magnifying the resulting displacement of one wall of the cell containing the gas. This detector is quite comparable in sensitivity with a Hilger–Schwarz thermocouple and can be made faster in response time (down to a few milliseconds). It has one great advantage over the thermocouple and bolometer, namely the receiver area is quite large; its disadvantages are that it is not quite so robust and can be damaged by overloading.

An entirely different type of thermodetector is the evaporograph first successfully developed by Czerny[7] in 1929. A crude form of this type of detector had been made by Sir John Herschel (the son of Sir William) in 1840.[8] In Czerny's form of the evaporograph, the infra-red radiation causes oil to evaporate from those portions of a thin film on which the radiation falls. The resulting change in film thickness is made visible by an interference technique. In sensitivity and speed, this thermodetector is considerably inferior to a thermocouple or bolometer, but it is a simple and indeed the only direct way of producing a picture by infra-red radiation of intermediate wavelength ($5–20\,\mu$). All other methods for this region involve scanning in some form or other.

Photodetectors

Until well into the twentieth century, the only photodetector for infra-red radiation was a specially sensitized photographic plate, which could cover only a minute range, namely from the

ge of the visible ($0.75\,\mu$) to $1.0\,\mu$. Although the short range
the photographic plate greatly limits its applicability, the
success of infra-red sensitive emulsions in the penetration of a
light haze and of camouflage must be well known to all of
you. Less well known perhaps is its application to the scientific
side of art history.[9] Intensive research has still failed to
extend this range beyond $1.3\,\mu$, but the Second World War
stimulated research into other means of photodetection which
had begun to be uncovered shortly after the end of the First
World War. There is little doubt that the problem of detect-
ing a military target by the 'passive' method of observing the
heat radiated from it was primarily responsible for the recent
remarkable advances in the techniques of photodetection of
infra-red radiation. Yet the applications of these new photo-
detectors have been of far greater importance in pure and
industrial science than in the military field.

The possibility of detecting infra-red radiation through its
property of quenching or stimulating phosphorescence, pre-
viously excited by visible or ultra-violet radiation, was not
investigated seriously until nearly a hundred years after
Becquerel's discovery of this phenomenon about 1840. During
the last war, very neat detectors of this type were developed in
the United States from alkaline-earth sulphides (such as calcium
and strontium) containing traces of other elements such as
barium and samarium.[10] Such detectors could be activated by
a radioactive button, and they were often invaluable after dark
as a means of rallying paratroopers to a leader, who only had
to carry an ordinary electric torch fitted with an appropriate
infra-red filter which removed all visible light from the beam.
Unfortunately, the sensitivity of the phosphor type of detector
is rather poor. As a consequence, it has a limited field of
application, but it possesses the great merits of compactness
and simplicity of operation.

A picture-forming photodetector which covers much the same
range as a photographic plate or infra-red phosphor, but is more
convenient for many purposes than the former and considerably
more sensitive than the latter, is the photo-emissive detector.
This is often referred to as an image converter, since it converts
an infra-red image of an object into a visible picture. This is
accomplished by forming the image on a film of caesium and
silver oxides. The work function of this material is sufficiently
low for the energy of the infra-red photon to cause emission of
the electron on which it falls. These electrons are then accelerated
and focused (by suitable electric fields) on to a screen of wille-
mite, so producing a visible fluorescent image.[10] Such detectors
are widely used during the last war. For instance, tanks and
other vehicles could be clandestinely driven at night by fitting
them with infra-red headlights, the driver being provided with
suitable headgear incorporating an image converter. Another
obvious military application was to sniping. Since the war,
detectors of this type have been found very valuable in astro-
nomical and other areas of purely scientific research.

The remaining three types of photodetectors listed in Table 2
differ from the three just discussed in that no direct image of the
infra-red source is produced. By far the most important are the
photoconductors, the first of which was developed by Case,[11]
who, in 1920, found that infra-red photoconductivity in thallium
sulphide could be made the basis of a detector sensitive to about
$0.9\,\mu$. Such Thalofide detectors were used by the Afrika Korps
to operate a *Lichtsprecher*, i.e. to talk along an infra-red beam.
During the war, the Germans and, somewhat later, the British and
the Americans, developed a series of infra-red photo-conductors
which penetrated much farther into the infra-red, namely,

Lead sulphide with a limit near $3.3\,\mu$ (Fig. 3).
Lead selenide with a limit near $8.0\,\mu$.
Lead telluride with a limit near $6.0\,\mu$ (Fig. 4).
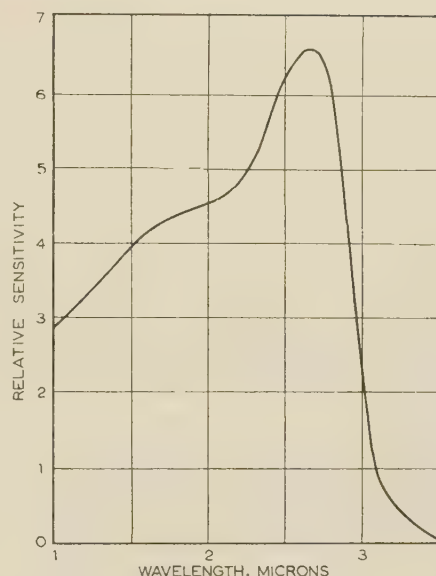
Fig. 3.—Sensitivity of a lead-sulphide detector to infra-red radiation
between 1 and $3\,\mu$.

This detector operates at room temperature. The range of sensitivity can be
extended to about $3.3\,\mu$ by going to liquid-air temperature.
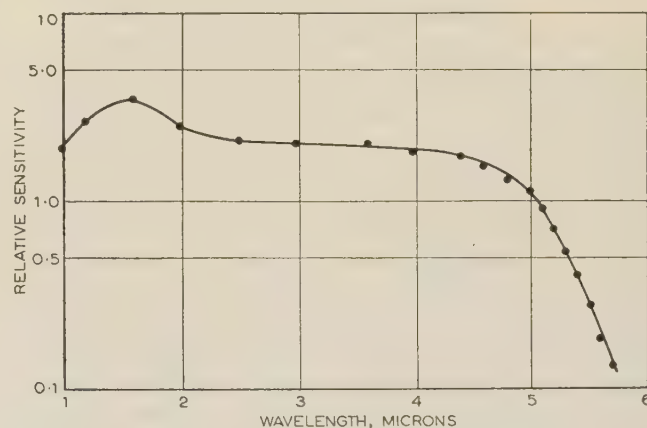
Fig. 4.—Sensitivity of a lead-telluride detector to infra-red radiation
between 1 and $6\,\mu$ at $90°\,$K.

This detector cannot be operated at room temperature.

These detectors have revolutionized work in this region of the
spectrum because the best are roughly a hundred times more
sensitive than any existing thermodetector and can be over a
hundred times faster in their response time, i.e. less than 10^{-4} sec.
Lead-sulphide and lead-telluride cells are now readily available
from commercial firms, but the selenide cells have not yet
progressed beyond the laboratory stage. The manufacture of
all these cells is a mixture of art and science, and those with the
best characteristics still have to be selected from large batches
with a considerably lower performance. Nevertheless, within
its spectral range the average lead-sulphide or lead-telluride cell
is still far superior to any thermocouple or bolometer.

During the past few years, some other photoconductive cells
have been developed. One of the most promising is indium
antimonide,[12] which is sensitive to about $7\,\mu$. Unlike lead
telluride or lead selenide, it does not require to be cooled to
liquid-air temperature, but operates at room temperature.
Other advantages are that the sensitive layer does not require
to be kept *in vacuo* and that the cell has a resistance of a few

hundred ohms, in contrast to several megohms for a good lead-telluride cell. Most important of all is the fact that the time-constant is probably less than 10^{-7} sec. Unfortunately, the sensitivity of these cells appears to be so much less than that of a good thermocouple that their field of application is at present limited to uses where speed of response is the dominant requirement.

One of the outstanding problems in the detector field is the provision of a sensitive, fast photodetector which responds to wavelengths as long as 20 or $30\,\mu$. Recent investigations in Britain[13] and the United States[14] have shown that silicon may fulfil this requirement when used at liquid-helium temperatures.

Photo-voltaic detectors have been made from thallium sulphide and lead sulphide, but their sensitivity was inferior to the corresponding photoconductive cell and this line of research has not been pursued.

More interesting is the development of a photo-electromagnetic detector based on indium antimonide.[15] Infra-red radiation incident on a single crystal of indium antimonide generates hole-electron pairs near the surface; these diffuse inwards and, by having the crystal between the poles of a magnet, the holes and electrons drift away from one another to produce a voltage between the two ends of the crystal. The response time is given as 10^{-6} sec or less, and the equivalent noise input near $7\,\mu$ as 10^{-9} watt, implying a sensitivity not much less than that of an average thermocouple.[16]

INFRA-RED SPECTROSCOPY

In describing the development of infra-red detectors, mention has been made here and there of applications, especially to military problems, since in the first instance the development was often stimulated by a military requirement. However, by far the most important use of infra-red detectors has been in infra-red spectroscopy, which now occupies a place in science comparable with X-ray analysis as a means of investigating the structure of matter. Whereas X-ray analysis is based on the fact that the regular arrangement of the atoms in a crystal gives rise to an X-ray diffraction pattern from which the positions and inter-nuclear distances of the atoms can be derived, infra-red analysis is based on the fact that atoms are not rigidly fixed but vibrate about mean positions and the consequent vibration 'pattern' gives rise to a characteristic infra-red absorption spectrum which can be analysed to yield information on inter-atomic distances, inter-atomic forces and molecular configurations. The principal advantage which infra-red analysis has over X-ray analysis is that it is much more powerful when matter is in the gaseous or liquid state and when hydrogen atoms have to be located. X-ray analysis is much more powerful for the crystalline state and especially for metals, where infra-red methods fail almost completely.

Some of the basic principles of infra-red analysis can be illustrated by considering a simple example such as the carbon dioxide molecule, in which the atoms have the symmetrical linear arrangement shown in Fig. 5. This molecule has three fundamental modes of vibration, from which every other possible mode of vibration can be built. Whether any of the funda-mental or other modes can be observed in the infra-red region depends on whether, associated with that mode, there is a periodic change in the dipole moment of the molecule. If so, then the molecule can absorb (and emit) infra-red radiation of the corresponding frequency. For instance, the unsymmetrical mode of vibration illustrated in the uppermost part of Fig. 5 will produce such a change in dipole moment, since the charge on either oxygen atom differs from that on the central carbon atom. On the other hand, the middle mode of vibration being quite symmetrical does not change the permanent dipole moment of

WAVELENGTH, CM
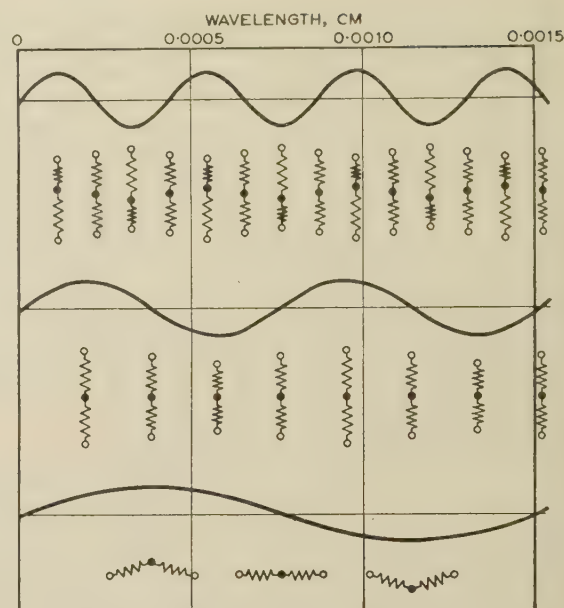
Fig. 5.—The three normal modes of vibration of the carbon-dioxide molecule.

● Carbon atom.
○ Oxygen atom.

the molecule (which happens to be zero in this case) and is therefore called an 'inactive' fundamental, i.e. unobservable in infra-red absorption. The third fundamental will be 'active', since it gives rise to a change of dipole moment perpendicular to the line of the molecule.

If the arrangement of the three atoms had been OOC instead of OCO, or if the three atoms had been arranged in an isosceles triangle instead of in a line, all three fundamentals would have been active. Thus, by observing how many of the fundamental modes of vibration appear in the infra-red spectrum of a mole-cule, it is possible to deduce the configuration of the atoms. In this way it has been possible to establish the structure of a large number of simple molecules, provided they possess at least one element of symmetry, although the method is, of course, most powerful for molecules having several symmetry elements.

The fundamental absorption frequencies of carbon dioxide which we have just discussed occur near $4 \cdot 25\,\mu$ (top mode) and $14\,\mu$ (lowest mode). If either of these absorption bands is studied in detail, it is found to consist of a series of absorption lines. These arise from the fact that the molecule is rotating as well as vibrating, and this rotational motion is also 'quantized' into a discrete series of rotational frequencies which are superimposed on the much higher vibrational frequency. By measuring the separations between such rotational lines, it is possible to deduce the moment of inertia of the molecule and hence the distance between the carbon and oxygen atoms. Clearly, such a method of determining inter-atomic distances has considerable limita-tions, since no molecule has more than three moments of inertia, but as soon as we tackle more complex and more unsymmetrical molecules the number of independent inter-atomic distances exceeds three quite quickly. By using isotopic substitution, it is possible to extend the scope of this method, and within its range it is by far the most accurate available to the molecular physicist. For instance, the separation between the oxygen and hydrogen atoms in water is now known to an accuracy of $0 \cdot 001$ Å, or 10^{-11} cm $(0 \cdot 958$ Å$)$, while the angle between the two OH bands can be given to within one minute of angle $(104° \ 27')$.

It might appear from the foregoing that structural analysis by infra-red can only be applied to a relatively small number of molecules in the gaseous state. While this is true of the very accurate techniques outlined above, there are less refined methods for which the rotational fine structure is not required, and these can be applied to molecules in the liquid or solid state. The 'selection rules' which determine whether a particular mode of vibration will be active in absorption are less strictly obeyed in the condensed state. Here, forbidden frequencies may be weakly active but can usually still be recognized. Furthermore, the majority of interesting large molecules are built up from units (such as CH_2, CH_3, NH_2, $C=O$, $C=C$, $C\equiv N$, $C\equiv C$ groups) which preserve their vibrational frequencies almost unchanged in any environment. The characteristic fundamental vibrational frequencies of the majority of these important molecular units have now been identified, and this frequently makes it possible to answer certain structural questions which the chemist cannot readily answer by more traditional methods. Since every species of molecule must have its own vibration 'pattern', one of the most important features of the infra-red spectrum of a molecule is its uniqueness. As it is now possible to obtain the vital characteristics of the infra-red spectrum of almost any molecule in a few minutes, the infra-red spectrum is even the physical property by which a molecule can most readily be 'finger-printed'. The consequences in chemical analysis, whether in routine, control or research work, are very far-reaching. No modern chemical laboratory can afford to be without at least one infra-red spectrometer, and the larger ones have several in continuous use. One of the leading chemists in the United States, Professor R. B. Woodward of Harvard University, has recently expressed his opinion in the following words:[17]

. . . But no single tool has had a more dramatic impact upon organic chemistry than infrared measurements. The development, just after the second Great War, of sturdy and simply operated machines for the determination of infrared spectra has permitted a degree of immediate and continuous analytical and structural control in synthetic organic work which was literally unimaginable fifteen years ago. The power of the method grows with each day, and further progress may be expected for a long time to come.

Besides being a very powerful analytical tool, infra-red spectroscopy provides the chemist with data from which the entropy and other thermodynamic functions of a molecule can be computed. Some of this information, which is vital to chemical industry, would otherwise be unobtainable.

The vibrational frequencies of a molecule depend on two factors, the masses of the atoms, and the forces between them which are brought into play when the inter-nuclear distances are changed. Consequently, infra-red spectra provide almost the only and certainly the most direct source of information about intra-molecular forces. We now know the force constants associated with the stretching of all the common chemical bonds and the deformation of the angles between them. This fascinating subject would require a whole lecture to itself and is so technical that I can do no more than refer to it here. Instead, I propose to discuss some of the applications of infra-red spectroscopy which may be of more general interest.

APPLICATIONS OF INFRA-RED SPECTROSCOPY

The Ammonia Clock

The first detailed examination of the infra-red spectrum of ammonia gas was made in 1928 by the late Sir Robert Robertson[18] at the Government Chemist's Laboratory across the way in St. Clement's Inn Passage. He noticed a curious doubling effect in some of the bands. This was more fully investigated at the
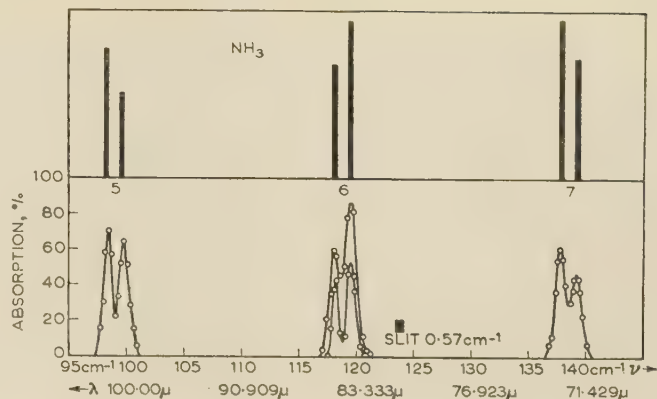
Fig. 6.—The doubling phenomenon in the rotational spectrum of the ammonia molecule.

The lower half of the diagram shows the observed spectrum between 70 and 105 μ; the upper half shows the predicted spectrum.
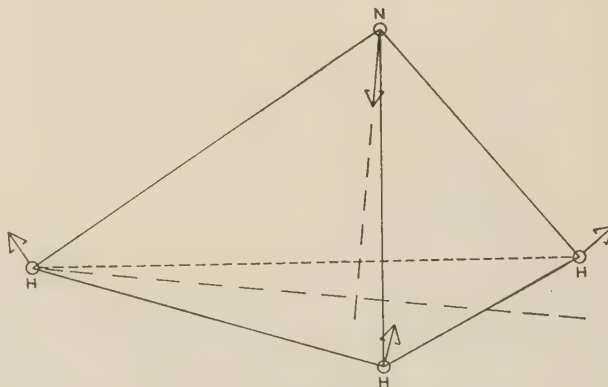
Fig. 7.—Pyramidal structure of the ammonia molecule showing one of the normal modes of vibration.

University of Michigan, where it was shown that the lines in the pure rotation spectrum were all doublets instead of single lines (Fig. 6). The explanation of this effect is that the ammonia molecule is a rather flat pyramid (Fig. 7) and the intra-molecular forces are such that it is relatively easy for the nitrogen atom to pass through the plane of the hydrogens, i.e. for the molecule to 'turn itself inside out'. This causes every energy level to be split into two separate levels and the whole spectrum is effectively doubled. From the separation of such doublets, Dennison and Uhlenbeck[19] were able to deduce the inversion frequency of the ammonia molecule and to predict that a corresponding absorption band should be observable in the microwave region close to 1 cm. In 1933, Cleeton and Williams[20] carried out a pioneering experiment in microwave spectroscopy which confirmed this prediction. During the war, great advances took place in the production and detection of centimetric microwaves, making it relatively easy to work at such wavelengths, and shortly after the war it was realized that this characteristic inversion frequency of the ammonia molecule could be used as a frequency standard. An ammonia 'clock' was constructed at the U.S. Bureau of Standards[21] which was accurate to about 1 part in 10^7. Although this has now been superseded by a caesium 'clock' (recently perfected by Essen and Parry[22] at the National Physical Laboratory to an accuracy of 2 parts in 10^{10}), the origin of these new methods of measuring time lies in the curiosity of some scientists and their desire to understand a small detail in the fine structure of an infra-red absorption spectrum.

The Water-Vapour Line at 1·25 cm

Another example of this same completely unpredictable use of a piece of pure research comes from the water-vapour spectrum. Although the water molecule is one of the simplest triatomic molecules, the rotational fine structure of its infra-red spectrum is extremely irregular and complex because it is an asymmetrical rotator, i.e. the three moments of inertia are unequal. The elucidation of this fine structure took several years, and the only obvious reward was a more accurate knowledge of the structural constants of the water molecule. However, during the war, anomalies were noticed in the atmospheric transmission of microwaves if the wavelength was a little over 1 cm, and these appeared to be associated with the humidity. This was extremely puzzling, since the pure rotation spectrum of water vapour would not be expected to extend into the centimetric region. A

noticed that nitrous-oxide bands can be detected in the earth's atmosphere, but the amount of this laughing gas is too small to produce significant effects! Again, one may speculate on the character of our world if the concentration had happened to be somewhat higher. In addition to the gases just mentioned, methane and HDO have also been identified in the earth's atmosphere.

To obtain such a spectrum, the sun is used as the source of infra-red radiation. A careful study of the emission from various parts of the sun's surface has proved that carbon monoxide exists in the outer layers of the sun's atmosphere.

Even more interesting is the information which can be obtained about certain gases in the atmospheres of the planets from a study of the solar infra-red radiation reflected by them.[23] Some of these results are summarized in Table 3. Notice that the atmo-

Table 3

PLANETARY ATMOSPHERES

Amount (cm at N.T.P.)

	O_2	N_2	CO_2	CH_4	N_2O	O_3	NH_3
Earth	$1·68 \times 10^5$	$6·25 \times 10^5$	220	1·7	0·8	0·3	—
Venus	—	—	1×10^5	—	—	—	—
Mars	—	—	440	—	—	—	—
Jupiter	—	—	—	$1·5 \times 10^4$	—	—	700
Saturn	—	—	—	$3·5 \times 10^4$	—	—	—
Uranus	—	—	—	$1·5 \times 10^5$	—	—	—
Neptune	—	—	—	$2·5 \times 10^5$	—	—	—

careful re-examination of the very complex energy-level scheme of the water molecule revealed two levels, at $447·24 \, \text{cm}^{-1}$ and $446·50 \, \text{cm}^{-1}$, between which an active transition was allowed which would give rise to an absorption line of frequency $447·24 - 446·50 = 0·74 \, \text{cm}^{-1}$, or a wavelength of $1·25 \, \text{cm}$.

Atmospheric Absorption—Terrestrial and Planetary

The absorption of solar infra-red radiation by the atmosphere has been studied ever since 1840, when Sir John Herschel[8] found he could detect at least two bands with his primitive evaporation thermograph. Herschel was only able to reach about $1·2 \, \mu$; Langley's bolometer enabled him to reach $5 \, \mu$, and by 1900 the spectrum was extended (albeit crudely) to beyond $10 \, \mu$. A rough plot of such an atmospheric spectrum (under low dispersion) is shown in Fig. 8 as far as $15 \, \mu$. It will be seen that it is prin-

Fig. 8.—Absorption of solar infra-red radiation between $0·7$ and $15 \, \mu$ by the earth's atmosphere.

cipally due to water vapour and carbon dioxide. The two fundamentals referred to in our earlier discussion of the spectrum of carbon dioxide can be seen at $4·2 \, \mu$ and $15 \, \mu$. Of particular interest is the ozone band near $10 \, \mu$, since this is where the maximum occurs in the infra-red radiation emitted by the earth's surface at normal temperatures. Indeed, the whole radiation heat balance of the atmosphere is controlled by these absorption bands of water, ozone and carbon dioxide. If the shapes and force constants of these molecules had been different, we should have had an entirely different climate on this globe. It will be

sphere of Mars contains only about twice as much carbon dioxide as that of the earth, so that life on Mars could not be excluded on that account. However, the amount of carbon dioxide on Venus rules out all possibility of life on that planet. The atmospheres of the next four planets all contain vast quantities of methane, while that of Jupiter also has an appreciable concentration of ammonia. No explanation has yet been given of these facts, which must be accounted for in any satisfactory theory of the origin of the solar system.

Analysis of Petroleum[24]

One of the earliest and most successful practical applications of infra-red spectroscopy was to the analysis of petroleum. This was developed during the war to replace older, more tedious and less precise methods based on the determination of refractive indices and 'aniline points'. Petroleum consists of a mixture of hydrocarbons; each hydrocarbon has its characteristic skeleton of carbon and hydrogen atoms and that skeleton gives rise to a characteristic infra-red vibration spectrum. For instance, three possible 'octanes' are illustrated in Fig. 9. Each

Fig. 9.—Arrangement of the eight carbon atoms in three different octanes.

s eight carbon atoms, but the carbon atoms may be arranged a line (*n*-octane) or to give branched structures such as 2.4 trimethylpentane or 2.3.4 trimethylpentane. The anched hydrocarbons have very desirable properties in an ternal-combustion engine, whereas linear hydrocarbons have ther poor properties. The infra-red spectra of all octanes are tremely different. From Fig. 10 it will be seen that even the
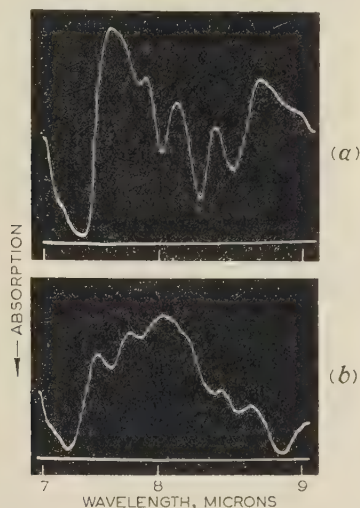
Fig. 10.—Infra-red absorption bands between 7 and 9 μ.

(*a*) 2.2.4 trimethylpentane.
(*b*) 2.3.4 trimethylpentane.

o branched octanes are easily differentiated from one another the bands occurring in quite a short range of the spectrum, mely 7–9 μ. The success of this method in the petroleum dustry was so great that it gave a tremendous stimulus first to ie commercial production of simple, reliable infra-red spectro-eters, and then to their application to every branch of the iemical industry. Infra-red analysis is now used extensively the large-scale production of petroleum, synthetic rubber, astics, drugs, antibiotics and synthetic fibres.

The Diamond Problem

The pioneer work of the late Sir Robert Robertson on nmonia[17] has already been mentioned. Another of his dis-veries was that diamonds may be divided into two classes n the basis of their infra-red absorption spectra.[25] Among veral hundred diamonds which he studied, he found about ilf a dozen which showed a spectrum quite different from that hibited by all the others. He called the common variety type I id the rare variety type II. He also established differences tween the two types in their ultra-violet absorption. These fferences are illustrated in Fig. 11. It will be noticed that pe II has no absorption peak beyond 5 μ, whereas type I is a complex of bands between 7 and 12 μ. In the ultra-violet, pe II diamonds are transparent down to about 2 300 Å, but pe I diamonds begin to show absorption near 3 000 Å. It was u id by later workers[26] that type I diamonds had peculiar tra streaks in their X-ray diffraction patterns which were not served in type II diamonds.

A great deal of subsequent work on this problem has still iled to produce a complete explanation. Instead it has shown at the average diamond is so far from being a perfect crystal at it requires a conference to itself each year! It now appears a the division into two classes was the first step in uncovering uch wider variation among diamonds. Although type II n be classed uniquely by the absence of the long-wave set
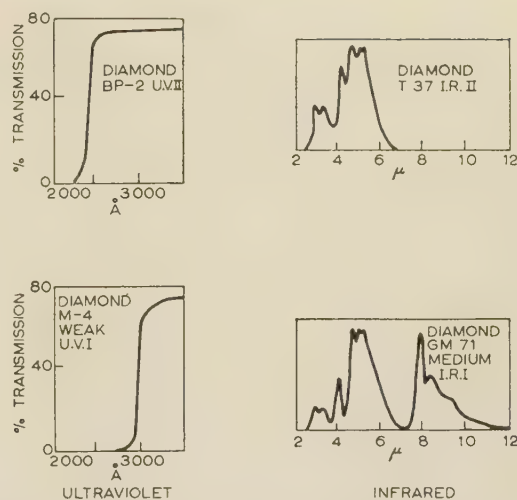
Fig. 11.—Spectrophotometric classification of diamonds.

The two curves on the left show the difference between the transmission of type I and type II diamonds in the ultra-violet. The two curves on the right show the difference between the absorption of type I and type II diamonds in the infra-red.

of bands, type I exhibit a virtually infinite variation in the overall and relative intensities of these bands. It is more correct to say that diamonds vary between extreme type I to type II, since the intensity of the bands beyond 6 μ varies over an extremely wide range. The most likely explanation is that the type II diamond comes closest to being a perfect crystal, since its infra-red and ultra-violet spectra and X-ray diffraction pattern all fit closely the theoretical predictions, whereas type I diamonds show anomalies in each of these properties.[27] The spectro-scopic anomalies can be accounted for formally by the presence of impurity atoms or impurity centres (e.g. missing carbon atoms). Probably both causes are operative, since diamonds are known to contain chemical impurities, but the quantities observed are rather small to explain the intensity of the effects.

A most interesting sub-class of type II diamond has recently been discovered, now known as type IIb. These were first noticed by their infra-red spectra,[28] which showed some extra lines in the region of 3 μ. However, the really remarkable property was discovered by Custers,[29] who found that they were semi-conductors. When put between electrodes with a potential of a few volts, they transmit a current which builds up so rapidly that the diamond soon becomes red hot.

This example was chosen to illustrate the power of infra-red analysis to open up a new field of research. It is often so sensitive to small variations in structure that it will spot an anomaly in structure more efficiently than any other method. Unless the irregularities in a structure are considerable, they can easily be missed by X-ray methods. Infra-red studies are likely to become of considerable importance in research in the non-metallic part of the solid-state field and have already yielded much valuable information about the electronic energy levels in silicon and germanium.[30]

Structure of Biologically Important Molecules

By far the most interesting molecules occurring in nature are those which have been evolved for a specific biological purpose, and of these the protein molecule and the nucleic-acid molecule are of fundamental importance in the maintenance and propaga-tion of every form of life. For instance, protein molecules are either the main or the most essential constituents of blood, muscle, tendons, skin, hair, feather, enzymes, nerve fibres and the anti-bodies which are unconsciously synthesized by an

animal to protect it from various diseases. Nucleic-acid molecules are the most essential constituents of cell nuclei. There are two forms of nucleic acid, known as deoxyribonucleic acid and ribonucleic acid. The former has been proved to be intimately connected with the transmission of genetic characteristics from one generation to the next, while ribonucleic acid appears to be essential to the synthesis of proteins in the growing cell. However, the way in which these molecules perform their biological tasks is still a complete mystery. We cannot hope to understand this mystery unless we know the structure of these molecules. The biochemist has made great progress in showing that each of these molecules is a polymer and in identifying the repeat unit. However, the configuration of the molecule is a problem for the physicist, and various lines of evidence indicate strongly that the configuration of the molecule must play a dominant role in its mode of action. While the more highly developed methods of X-ray analysis have so far played the major role in elucidating some of the main features of the structure of proteins and nucleic acid, infra-red analysis has given important independent evidence and is likely to have an increasingly important role in the future developments, since X-ray methods can only be applied to the crystalline or highly ordered parts of structures and are virtually useless in the liquid state.

Proteins

Chemists have proved that the protein molecule is built from polypeptide chains of the type shown in Fig. 12. However,

Fig. 13.—The polypeptide chain.

(a) Helical form.
(b) Fully extended form.

Fig. 12.—Diagrammatic representation of a polypeptide chain.

The repeat unit consists of a CO, an NH and a CHR group, where R differs according to the amino-acid from which the unit is derived. Proteins may contain over ten different amino-acid residues.

physical methods are required to establish the spatial configuration of the chain and the spatial relationships of these chains to one another in each protein molecule. By studying synthetic polypeptides (made from a single amino acid) it has now been shown, by a combination of X-ray and infra-red analysis, that these can exist either in the fully extended form (known as the β-configuration) or in a coiled helix (known as the α-form). This is roughly illustrated in Fig. 13. It will be noticed that in the α-form, the CO and NH bonds are parallel to the molecular axis, whereas in the β-form these bonds are

perpendicular to direction of alignment of the chain as a whole. The CO and NH bonds have characteristic absorption bands (near $6 \cdot 0$ and $3 \cdot 0 \mu$ respectively). Using polarized infra-red radiation, it is found that the β-form gives strong absorption in these regions only when the plane of polarization is at right angles to the direction of orientation of the molecular chains. For the α-form, the plane of polarization has to be parallel to that direction.

When fibrous proteins such as silk, hair or porcupine quill are studied with polarized infra-red radiation, it is found that silk closely resembles the β-polypeptide, while hair and porcupine quill resemble the α-polypeptide.[31] For the globular proteins such as haemoglobin, the problem is much more difficult because the polypeptide chains are not now all oriented in the same direction, and a great deal of work has still to be done before the structure of the protein part of haemoglobin is established.

Deoxyribonucleic Acid (DNA)

This is also a giant polymer molecule, shown by the chemist to consist of a repeating pattern of a sugar residue (ribose) linked to a phosphate group and a base, in which the key group is a hexagon of four carbon and two nitrogen atoms. Certain absorption bands in the infra-red spectrum of DNA can be assigned to vibrations in the plane of this hexagon. When threads of nucleic acid are examined in polarized infra-red radiation, it can be seen that the bases (hexagons) must be perpendicular to the axis of the polymer.[32] From X-ray data, Crick and Watson[33] have shown that the nucleic-acid molecule consists of two polymer chains wound spirally about a common axis and linked together through the bases. This is illustrated in Fig. 14, where the bases appear on the left-hand view as lines and the sugar residues are the pentagons appearing at regular intervals along the helices.

Crick and Watson have proposed that the way in which the bases are linked together is quite specific, namely as shown in Figs. 15 and 16. If this hypothesis is accepted, it is possible to see, in principle, how nucleic-acid molecules could replicate and thus how the information which determines the genetic characteristics of a species can be transferred from one generation to the next. The basic idea can be appreciated from Fig. 17. On the right is an intact DNA molecule in which the letters A, T,

Fig. 14.—Two photographs of a model of the DNA molecule due to Crick and Watson.[33]

The chemical bonds in the phosphate-sugar backbone are represented by wire. The pairs of bases are represented by metal plates. The fibre axis is represented by the (white) Perspex rod.
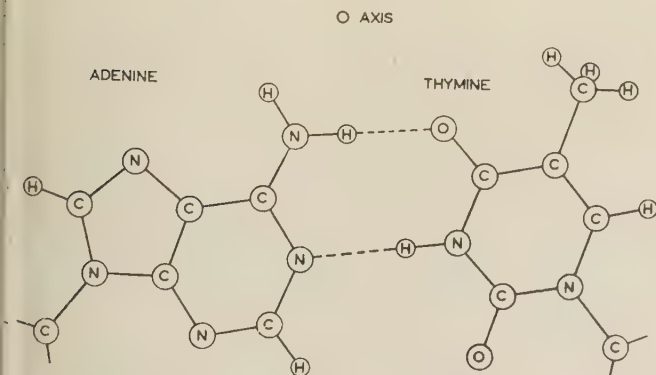
Fig. 15.—The pairing of the two bases adenine and thymine.
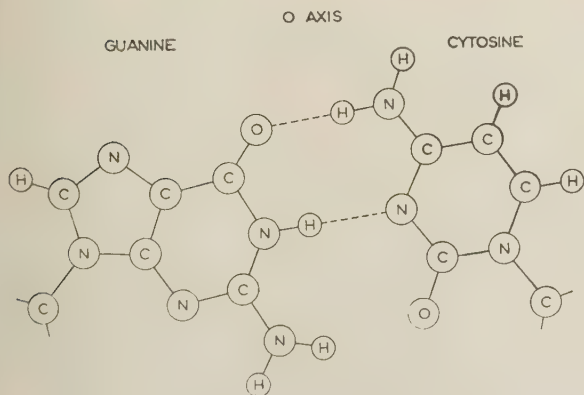
Fig. 16.—The pairing of the two bases guanine and cytosine.
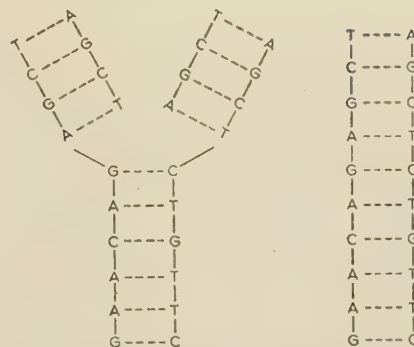
Fig. 17.—Possible replication mechanism for DNA (deoxyribonucleic acid).

C, G represent the bases linking the two polymer chains (i.e. adenine, thymine, cytosine and guanine). Let us suppose that the genetic information is contained in the arrangement of the bases along one chain, since it can be shown that the molecule is large enough for such a four-letter code to contain the required amount of information. Now suppose that when a cell divides to form two new cells, the DNA molecule splits as shown on the left side of Fig. 17. Each half of the original molecule can now act as a template on which the other half of the new molecule will form in a unique manner, since A can only bond to T and C to G.

It must be emphasized that this is only a hypothesis, but it is a great advance on anything which preceded it. We are approaching the stage when the way in which biological systems store and transmit information can begin to be understood in terms of a molecular model. When one considers the size, intricacy and clumsiness of the computing machine and how little it can do in the way of controlling any but the simplest operations, it should be clear that the communication and control engineer of the future may have a great deal to learn from the biophysicist who is probing around now to try and understand the elegant and compact way in which nature uses molecules to accomplish similar objectives. Infra-red analysis is playing a role in this key problem of biophysics.

Herschel indeed made a prophetic remark when he started a paragraph in one of his papers with the sentence:

Lastly, on the transmission of invisible terrestrial radiation—the subject which appears most pregnant with useful inferences for the common purposes of life. . . .

REFERENCES

(1) HERSCHEL, SIR W.: *Philosophical Transactions of the Royal Society*, 1800, **90**, pp. 284 and 437.

(2) MELLONI, M.: *Poggendorffs Annalen*, 1833, **28**, p. 371; *ibid.*, 1835, **35**, pp. 112, 277 and 385.

(3) LANGLEY, S. P.: *Proceedings of the American Academy of Arts and Sciences*, 1881, **16**, p. 342.

(4) COBLENTZ, W. W.: *Publications of the Carnegie Institution*, Washington, 1905, 1906 and 1908.

(5) ANDREWS, D. H., MILTON, R. M., and DE SORBO, W.: *Journal of the Optical Society of America*, 1946, **36**, p. 518.

(6) GOLAY, M.: *Review of Scientific Instruments*, 1947, **18**, p. 357.

(7) CZERNY, M., and MOLLET, P.: *Zeitschrift für Physik*, 1937, **108**, p. 85.

(8) HERSCHEL, SIR J.: *Philosophical Transactions of the Royal Society*, 1840, **130**, p. 1.

(9) Report of the National Gallery, 1956.

(10) SUTHERLAND, G. B. B. M., and LEE, E.: *Reports of the Physical Society on Progress in Physics*, 1948, **11**, p. 144.

(11) CASE, T. W.: *Physical Review*, 1920, **15**, p. 289.

(12) AVERY, D. G., GOODWIN, D. W., and RENNIE, A. E.: *Journal of Scientific Instruments*, 1957, **34**, p. 394.

(13) ROLLIN, B. V., and SIMMONS, E. L.: *Proceedings of the Physical Society*, B, 1953, **66**, p. 162.

(14) BURSTEIN, E., OBERLY, J. J., and DAVISSON, J. W.: *Physical Review*, 1953, **89**, p. 331.

(15) HILSUM, C., and ROSS, I. M.: *Nature*, 1957, **179**, p. 146.

(16) SMITH, R. A., JONES, F. E., and CHASMAR, R. P.: 'Infra-Red Radiation' (Oxford University Press, 1957). A comprehensive and scholarly treatment of the problems of infra-red detection.

(17) WOODWARD, R. B.: in 'Perspectives in Organic Chemistry', edited by Sir Alexander Todd, 1956, pp. 155–184.

(18) ROBERTSON, SIR ROBERT, and FOX, J. J.: *Proceedings of the Royal Society*, A, 1928, **120**, p. 128.

(19) DENNISON, D. M., and UHLENBECK, G. E.: *Physical Review*, 1932, **41**, p. 313.

(20) CLEETON, C. E., and WILLIAMS, N. H.: *ibid.*, 1934, **45**, p. 234.

(21) LYONS, H.: 'Spectral Lines as Frequency Standards', *Annals of the New York Academy of Sciences*, 1952, **55**, pp. 831–871.

(22) ESSEN, L., and PARRY, J. V. L.: 'The Caesium Resonator as a Standard of Frequency and Time', *Philosophical Transactions of the Royal Society*, A, 1957, **250**, pp. 45–69.

(23) KUIPER, G. P.: *Reports of the Physical Society on Progress in Physics*, 1950, **13**, p. 251.

(24) SUTHERLAND, G. B. B. M., and SHEPPARD, N.: in "The Science of Petroleum', 1955 (Oxford University Press), **5**, p. 317.

(25) ROBERTSON, SIR ROBERT, FOX, J. J., and MARTIN, A. E.: *Philosophical Transactions of the Royal Society*, A, 1934, **232**, p. 463.

(26) RAMAN, SIR C. V., and NIKALANTAN, P.: *Proceedings of the Indian Academy of Sciences*, 1940, **11**, p. 379. Also LONSDALE, K., and SMITH, H.: *Nature*, 1941, **148**, p. 112.

(27) SUTHERLAND, G. B. B. M., BLACKWELL, D. E., and SIMERAL, W. G.: *Nature*, 1954, **174**, p. 901.

(28) BLACKWELL, D. E., and SUTHERLAND, G. B. B. M.: *Journal of Chemical Physics*, 1949, **46**, p. 9.

(29) CUSTERS, J. F. H.: *Physica*, 1954, **20**, p. 183. Also *Nature*, 1955, **176**, pp. 176 and 360.

(30) FAN, H. Y.: *Reports of the Physical Society on Progress in Physics*, 1956, **19**, p. 107.

(31) AMBROSE, E. J., and ELLIOTT, A.: *Proceedings of the Royal Society*, A, 1951, **206**, p. 206.

(32) SUTHERLAND, G. B. B. M., and TSUBOI, M.: *ibid.*, 1957, **239**, p. 446.

(33) CRICK, F. H. C., and WATSON, J. D.: *ibid.*, 1954, **223**, p. 80.

DISCUSSION ON

'TRANSISTOR CIRCUITS AND APPLICATIONS'*

SOUTH MIDLAND RADIO AND MEASUREMENT GROUP, AT BIRMINGHAM, 20TH JANUARY, 1958

Mr. J. S. Roebuck: It often used to be stated that transistors operate satisfactorily in audio- and video-frequency amplifiers on the test bench, but when produced under mass-production conditions, the variation between batches is so great that it is impossible to obtain the same results. Is this still true?

Mr. J. F. Evans: What is the simplest method of achieving amplitude stability in a transistor oscillator over a wide temperature range? We have found it necessary to include a squaring circuit in the feedback loop of an oscillator, and to rely on a high-Q-factor circuit to remove the harmonics. In this way it seems possible to use germanium transistors from -50° C to 60° C, with small amplitude variation and low second-harmonic distortion, as compared with other means of stabilization which we have tried.

Secondly, can thyratron-type servo amplifiers using transistors as switches be operated with a partly reactive load?

Mr. M. H. Roberts: I have done experiments on pulse generation and found it possible to pass currents of between 70 and 100 amp through a transistor. Has the author any experience with such high currents?

With regard to the stabilization of circuits over a wide range of junction temperatures, I can find little mention of the fact that the base current has to be reversed in some cases, and manufacturers' curves give no information about this.

Another peculiar effect we have noticed is that, in switching off an inductive load, the voltage transient does not go as high as one would expect, and does not even reach the transistor turn-over voltage. There seems to be some mechanism by which

the energy is absorbed, and we think that the transient rise in voltage and dissipation causes the junction temperature very momentarily to be raised—sufficiently for the collector current to continue with zero base current and dissipate the energy in that way. In this way we have been able to switch off an inductive load without any protective measures other than the presence of the transistor itself.

In the square-wave oscillator, with a very simple circuit using only a transformer with feedback windings and resistors, is it necessarily the saturation of the core that causes the change-over, or can it be the result of the decrease in current gain as the current rises?

Dr. A. G. Milnes (*in reply*): The difficulties mentioned by Mr. Roebuck tend to occur if transistors are used to the limit of their performance, without ample design margins.

A conventional method of achieving amplitude stability in transistor oscillators is by using reference diodes to limit the swing, and a stability of 1% over a 40° C range has been obtained.* Mr. Evans's proposal should be of interest to those concerned with this problem over wide temperature limits.

Mr. Roberts's comments are interesting, but with switched-type operation into inductive loads it seems desirable that induced voltage surges be limited by capacitance or rectifier paths in parallel with the load. At the I.R.E.–A.I.E.E. Semiconductor Conference in Philadelphia in February, 1958, Dr. L. Bright reported some evidence of progressive deterioration in the transistors of switching convertor circuits unless switching-voltage spikes were suppressed.

* MILNES, A. G.: Paper No. 2368, May, 1957 (see **104 B**, p. 565).

* KRETZMER, E. R.: *Proceedings of the Institute of Radio Engineers*, 1954, **42**, p. 391.

A FLYING-SPOT FILM SCANNER FOR COLOUR TELEVISION

By H. E. HOLMAN, Associate Member, G. C. NEWTON and S. F. QUINN, Associate Member.

SUMMARY

Film moving with uniform velocity is scanned by a series of dis-
...ced rasters in such sequence that the system is applicable to 50 or
.../s conditions. Three photo-multipliers provide colour analysis of
. image, element by element, and directly produce a video-frequency
...al, so avoiding any necessity for accurate optical registration. A
...ticular equipment is described.

...he choice of RGB co-ordinates is examined, and arrangements are
...cussed to minimize colour flicker, to provide uniform field illumina-
... and to offer the most accurate picture reproduction.

(1) INTRODUCTION

.t has been demonstrated over a number of years that mono-
...ne television signals of the highest quality may be derived
...m continuously-moving film by the flying-spot method of
...nning. In this system the source of light appears on the face
.. a cathode-ray tube as a focused spot which is deflected to cover
. rectangular raster. An optical image of this raster fully
...cupies an aperture in the film gate in its width, or line direction,
...t is reduced in its height, or field direction. During each
...nning cycle, however, the height displacement of the spot
...age supplements the constant-speed motion of the film to
...onstitute thereon an explored field of normal aspect ratio and
...iform brightness. The spot intensity, modulated by passage
...ough varying densities of the film, is then converted directly
...o a video-frequency signal by the photo-multiplier tube.

...haracteristic of this system are the well-resolved, steady
...tures which exhibit a full range of contrast with no shading
...ors. A good signal/noise ratio is achieved, and reference
...ck is produced automatically to ensure a signal at uniform
...el. Because of these advantages it is logical to extend the
...hnique to colour television, where the flying-spot system is
...ique, since, with only one scanning process, no difficulties of
...cise colour registration are involved.

(2) CONSIDERATION OF BASIC REQUIREMENTS

The problem of generating television signals from motion-
...ture film can be approached from three main directions:

(*a*) An intermittent film projector used with a television camera.
(*b*) A mechanism where the continuous-motion of film is optically
compensated and which is employed either with camera or flying-spot
analysis.
(*c*) A continuous-motion flying-spot scanner using electronic com-
...pensation of the film movement.

...onventional projectors as employed in the cinema have
...ited application to television, the main obstacle being that the
...llisec available during field suppression is considerably
...ter than is required for reliable intermittent film transport.
...ecial 16 mm projectors with accelerated movements have
...ieved partial success, but the mechanical forces involved are
...ch that it may well prove impossible to apply this principle to
... larger 35 mm gauge. A further difficulty is that, unless there
...odification to the intermittent mechanism, the standard film

speed of 24 frames/sec is not suited to television field frequencies
of 50 frames/sec in Europe or 60 frames/sec in the United States.

Normal types of television camera tube require either constant
illumination, or else that a high-intensity image of the film be
produced by pulsed flashes during each field suppression period.
In the latter case, the memory characteristics of the tube then
permit the scan to be conducted in darkness but, unfortunately,
spurious signals are often generated throughout this interval,
and random patterns become superimposed upon the desired
image signal during scanning. This situation has been improved
by the recent introduction of reliable camera tubes of the vidicon
type, and under proper operating conditions, such tubes produce
no spurious signals and yet exhibit the full storage properties
necessary for the use of standard intermittent film projectors.
However, in common with all colour analysing systems which
use separate monochrome camera tubes, the overall definition
relies upon accurate registration of scanned patches produced
in three independent camera channels. This entails matching a
complete range of optical and electronic units, and presents
formidable problems both in the initial achievement and in
subsequent maintenance.

Group (*b*) includes projectors in which continuous motion of
the film is optically arrested by mirrors, or by a rotating prism.
The feature common to all these projectors is precise mechanical
compensation for film shrinkage which must be incorporated,
and they are therefore optically complex and necessarily expen-
sive. Additionally, since three camera tubes would be employed,
registration between individual channels imposes its further
limitation. While there have been examples of this type of
compensating mechanism applied to a flying-spot scanner, the
unfavourable optical efficiency places such equipment at a dis-
advantage; it is questionable, too, whether the complication of
optical immobilization is justified when allied to precise scanning
which can perform the same function.

The paper is concerned with group (*c*). The most successful
application hitherto has been with 50 field/sec monochrome
television systems, where an interlaced picture is produced from
a continuous-motion scanner operated at 25 frames/sec, the 4%
increase in speed being practically unnoticeable.[1, 2, 3] A twin-lens
system was used with a single scanning raster to form reduced
spot images which were modulated by the film, and alternately
selected by a rotating shutter to be fed into a photo-multiplier
tube. In the present case, which follows a later proposal, inter-
lace is effected by presenting a raster which electronically alter-
nates between two positions on the tube face and is imaged by
a single lens upon the moving film.[4] An extension of the same
principle covers a range of positions for the raster on the tube
face; by this means the conditions encountered in the 60-field/sec
television system may be satisfied.[5] With film travelling at
24 frames/sec, each successive pair of film frames is scanned with
five fields; two for the first, three for the next, then two, then
three, and so on. Above all, these methods impose stringent
requirements for accuracy, particularly in connection with elec-
trical deflection in cathode-ray-tube scanning, but they offer
better utilization of the tube phosphor, improved optical effi-
ciency, and important mechanical simplification.

The authors are with Electric and Musical Industries, Ltd.

(2.1) The Flying-Spot Colour System

In order to introduce the problems encountered and to present a coherent solution, a recently developed equipment will be described; it derives high-quality colour-television pictures from a wide range of commercial colour films. The apparatus has been designed for installation in a television station, and, if desired, may be used initially in a monochrome version to which colour facilities may later be added without structural change. Normally, it is not possible to effect the satisfactory conversion of a monochrome machine unless the design and specification of mechanical and optical components has been considered *ab initio* with relation to the ultimate colour application.

The spectral composition of light emitted by the phosphor is of prime importance in any flying-spot colour equipment; but it is sufficient at this juncture to establish that the scanning spot is presented in the form of a bright white raster upon the face of a cathode-ray tube. A specially-corrected high-aperture objective is used to produce an image of this spot upon the film
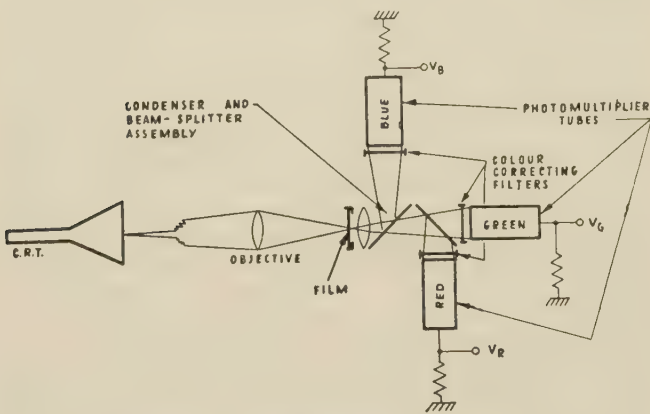
Fig. 1.—Flying-spot optical system.

(see Fig. 1), where certain components of white light are absorbed and the remaining coloured light is transmitted. A matched condenser system collects this coloured component and directs it through a beam splitter to produce uniformly illuminated patches in the correct proportions on three suitably-disposed photo-multipliers. By this method an instantaneous analysis of each element in the colour film is achieved in terms of the chosen primary colours.

(2.2) Immobilization of Film Motion

The sequence of patch positions to suit European conditions of 50 fields/sec (interlaced) with a film speed of 25 frames/sec is shown in Fig. 2. The scanning cycle starts when the leading edge of the downward-moving film frame reaches the horizontal centre-line of the gate. During $\frac{1}{50}$ sec from this instant the image of the spot traces the odd-line raster in an upward direction, i.e. opposite to the motion of the film. The vertical amplitude of the scanned patch is such that, in combination with the film travel, one complete frame is explored by the spot. The even-line raster follows, but is displaced in the image through a vertical distance equivalent to one-half the frame interval. Since this is precisely the distance that the leading edge of the film frame has descended during the $\frac{1}{50}$ sec period of the initial scan, the same film frame will be explored in an interlaced fashion during this following period
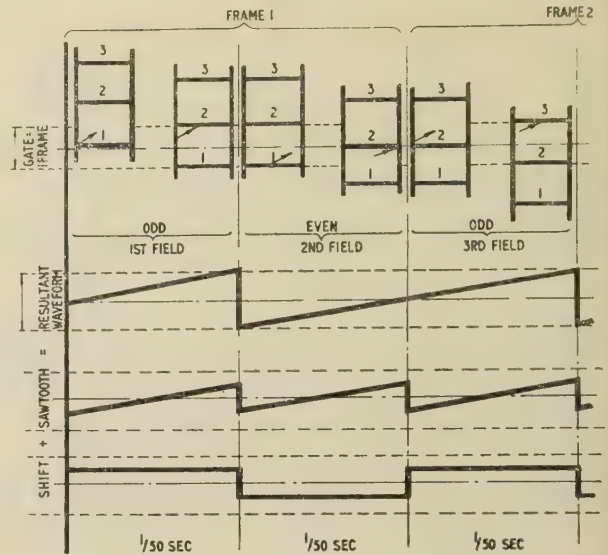
Fig. 2.—50 c/s scanning sequence.

of $\frac{1}{50}$ sec. At the end of this time the scanning cycle concludes with the spot image upon the trailing edge of the film frame near the horizontal centre of the gate. This procedure is subsequently repeated 25 times per second.

In the case of the 60 c/s television system, a film speed of 24 frames/sec is used, and the sequence of patch positions will be seen from Fig. 3. Starting the scanning cycle with the film frame-bar at one-tenth of the frame interval below the horizontal centre-line of the gate, the odd-line raster is completed in $\frac{1}{60}$ sec, during which time the leading edge of the first frame will have advanced two-fifths of the frame interval. The image of the even-line raster must therefore be presented at one-half the frame interval below the gate centre-line, and, after a further $\frac{1}{60}$ sec, will have completely scanned the first frame; this meanwhile has made a total advance of four-fifths of the frame interval. The second odd-line scan is therefore presented to the following film frame
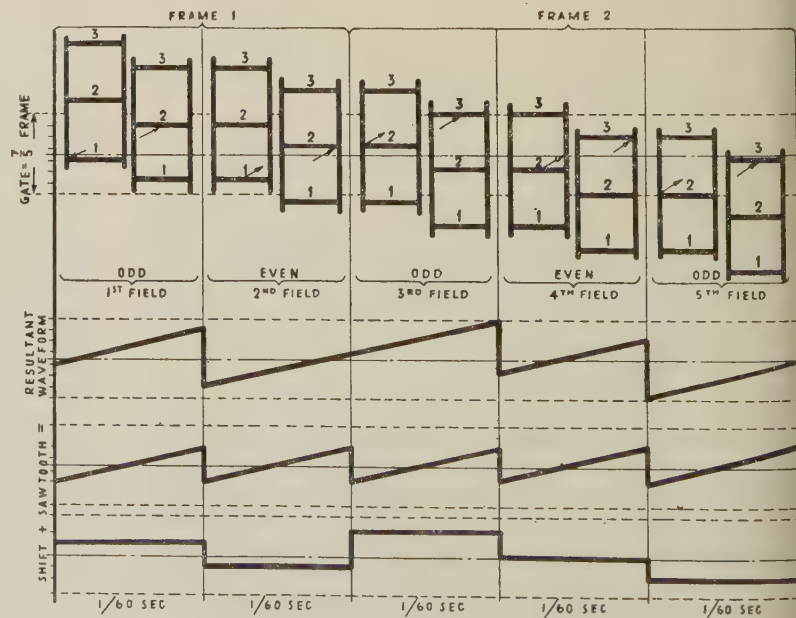
Fig. 3.—60 c/s scanning sequence.

a position one-fifth of a frame interval above the initial
starting point, and $\frac{1}{60}$ sec later the second even-line scan will
explore the same frame from below the gate centre-line. At the
end of this second complete scanning, i.e. the fourth field, the
whole area of the second frame is still exposed in the gate.
Therefore the next odd-line scan is able to start at the bottom
of the gate and complete its re-exploration of this frame, finishing
at the gate centre-line just $\frac{1}{12}$ sec from the start of the cycle.
The next five fields are completed in like manner, with the excep-
tion that the odd and even fields are interchanged. After this the
entire process repeats six times per second, every odd frame of
the film being scanned by two fields and every even frame by
three. It will be noticed that the total vertical excursion of the
scanning spot upon the tube face is at least 40% greater than in
the previous example, but that each raster is compressed in
height to approximately 60% of the normal aspect ratio. In
every $\frac{1}{6}$ sec, therefore, two rasters are presented in each of five
displaced positions upon the scanning tube, and, from the
resultant image analysis, the interlaced television picture is
reconstituted.

There are several essential requirements in order to produce
results of the highest quality. First, field pulses must occur at
strictly uniform intervals and rigorously in synchronism with the
film, whose velocity must remain constant at the rate of one-half
(alternatively two-fifths) frame per field exactly. Also, to avoid
distortion and defects of interlace, it is important to maintain
positional accuracy, i.e. any particular part of a film frame
must be presented to the corresponding points in two successive
scans at precisely the interval of field frequency.

In addition to the need for accurate scanning along the vertical
axis of the picture, the general geometrical characteristics of the
patch must be maintained irrespective of its position upon the
tube face. It is particularly important that a rectangular patch
is presented. For example, in extreme positions, it is required
to maintain interlace between the top corners of a patch generated
wholly in the upper half of the tube face, with corresponding
parts formed near the horizontal centre-line when the patch is
generated substantially in the lower half. This requires a
scanning field of excellent uniformity to avoid 'ragging' or dis-
continuity of vertical lines, particularly towards the side margins
of the interlaced field. In order to prevent trapezium distortion
of the image at the film gate, which may be produced optically
owing to angular displacement of the cathode-ray-tube face, any
optical and mechanical misalignment must be avoided. Replace-
ment cathode-ray tubes are therefore located in such a manner
that the screen is held normal to the axis defined by the optical
elements and the film aperture.

It will thus be appreciated that not the least of the major
development problems associated with this equipment have been
concerned with the design and construction of the film transport,
the scanning coils and the circuits necessary to maintain the
desired accuracy of picture immobilization.

(2.3) Arrangement of Scanners

A basic principle in the design of this equipment has
been the incorporation of preset mechanical and optical
assemblies. These reduce the number of operational adjust-
ments and confer flexibility to the layout without any sacrifice
of accuracy. For instance, the unit consists of a flat machined
base-casting 27 in from floor level, forming a free-standing
arrangement which allows access from all sides (Fig. 4). This
base-plate supports a single cathode-ray-tube assembly in main-
tained alignment with two inward-facing film-traction mech-
anisms. The mechanisms may be identical heads, either 16 or
35 mm, or a combination of one head for 16 mm and the other
for 35 mm film. Since their location is established from common
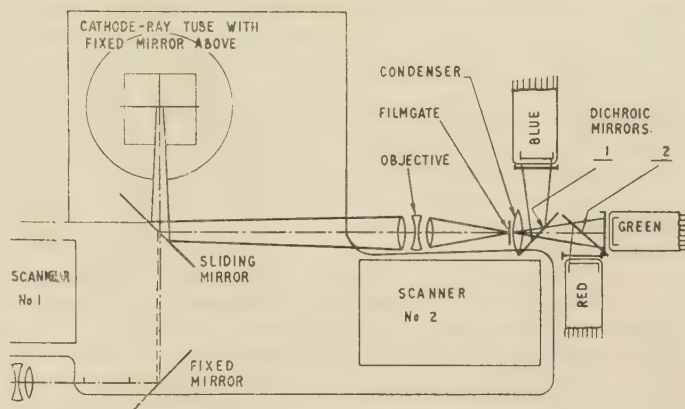
Fig. 4.—Arrangement of scanners and optical system.

datum faces, and since all projectors are of the same basic form,
they may later be interchanged as required (see Fig. 20).

The combined cathode-ray-tube support and mirror housing
serves the pair of projectors, and while each has its separate
colour analyser, they alternately receive illumination from the
common scanning tube. The cathode-ray tube is held face
uppermost against a horizontal location which is preset with
respect to the mirror assembly and the remainder of the system.
This orientation of the tube was chosen as the most desirable
for prolonged operating conditions since it reduces the possi-
bilities of screen contamination from spots, etc. Also, the
overall dimensions are reduced and the tube and scanning-coil
supports are rendered simpler than they would be if the tube axis
were horizontal. A reflector inclined above the tube face directs
the beam between the machines to strike one of a pair of vertical
mirrors which face the lenses of the respective projectors. The
mirror closer to the tube is arranged to slide vertically; in the
upper position it is clear of the beam, which is later reflected by
the farther mirror to No. 1 projector; in the lower position it
intercepts the beam and deflects it to the nearer projector No. 2.
The sliding mirror is moved by a remotely actuated solenoid and
provides an effective change-over between programme sources.
These mirrors, which have an anodized aluminium reflecting
surface deposited upon the optically worked flat surface, are, in
each case, supported by resting the front face itself against a
machined seating. Such an arrangement allows mirrors to be
replaced without adjustment, and simplifies the initial machining
since the only surfaces involved are accurately perpendicular to,
or at 45° with, the datum base. There are but two reflections,
each from plane surfaces; therefore no aberrations or distortions
are introduced in the multiplexing arrangements.

An image of the raster is formed at the 16 mm gate by a lens
of $2\frac{1}{2}$ in focus working at ×9·0 magnification, while for the
35 mm machine a lens of 4 in focus is used at ×4·2 magnification,
the initial object size being unaltered. These lenses, which have
a working aperture of $f2·0$, have been specially computed and
made for this exacting application. Besides the normal correc-
tions, particular care was taken to minimize vignetting, so that
a sensibly uniform illuminated field is produced, and to correct
for lateral chromatic aberration, so that images of different
colours formed in the same plane would be of identical size.
The latter defect has been corrected to such an extent that it is
impossible to distinguish any difference in magnification on a
reproduced white test chart as the colour components are
switched in or out. An important feature of this lens design
concerns the distribution of illumination across an elementary
image spot. In television applications which suffer linear

transformation rather than the logarithmic reproduction which occurs in photography, it is necessary to concentrate the transmitted energy into a uniformly bright spot with the minimum of low-intensity surround. This offers improved modulation by maintaining high contrast, with good blacks, at higher frequencies.

After transmission through the film, the beam of light is collected by a condensing system and analysed by presentation to three suitably colour-sensitive photo-multiplier tubes. One method of analysis is by means of partially reflecting mirrors with photocells sensitive in narrow spectral bands, either due to their own characteristics or in combination with suitable colour filters. This is wasteful because the partially reflecting film, e.g. aluminium, is absorbent, and also because useful light of all colours is reflected into each photo-multiplier tube, which is only permitted a narrow-band sensitivity. The preferred type of beam splitter consists of two dichroic reflectors which are themselves colour selective and of high efficiency, since the total of transmitted and reflected light is usually 95–98% of that incident. An elementary dichroic filter will reflect light of a given colour if a layer of material of higher refractive index, and one-quarter wavelength in thickness, is deposited upon the transparent base. The amount of reflection is then governed by the refractive-index differential, whereas the distribution of colour is a function of the optical path length offered by the layer. Therefore, an increase in the number of alternate layers traversed, as well as in the refractive-index difference, improves the spectral purity of the reflected colour. Even so, additional slight correction filters usually have to be applied to produce a true theoretical analysis.

A profound influence upon the uniformity of colour analysis is exerted by the condenser, whose function is to collect coloured light transmitted by the film and to image the whole aperture of the projection lens upon the face of the photo-multiplier tube. During its passage through the projector gate, every film frame is scanned in a number of positions, so that, after collection by the condenser, light reaches the photo-cathode by a number of different paths (see Fig. 5). Since the dichroic reflectors are positioned between these components, there is a variation in angle of incidence over successive fields amounting to some $\pm 15°$. The optical path length through a layer, and hence the spectral characteristic of the dichroic, depends upon the angle of incidence; and, owing to this change in characteristic, the transmission at different scan positions gives rise to different luminance values which result in colour flicker. It can be shown that the path difference is $2nt \cos r$,

where n = Refractive index of layer.
       t = Thickness of layer.
       r = Angle of refraction in the layer.

The difference in path length, therefore, is much less for small angles than for large, and a variation of 15° about normal incidence produces only about one-quarter of the difference that it would if the angle of incidence were 45°. It will be seen, therefore, that the effect of this variation over successive fields can be minimized by arranging the beam splitter to present its dichroic surfaces normal to the direction of the field scan and at 45° to the direction of the line scan, i.e. to distribute its light in a horizontal plane.[6] There may still be a variation in colour signal derived along a line which would impart a tint to the colour display, reddish on one side and greenish on the other. 'Colour tilt' controls are provided by which this condition is corrected; they separately vary the gains of the photo-multiplier tubes with a line sawtooth.[14]

There are, of course, certain values of the position and focal length of the condenser assembly which afford a minimum variation of incident angle, and their correct choice ensures
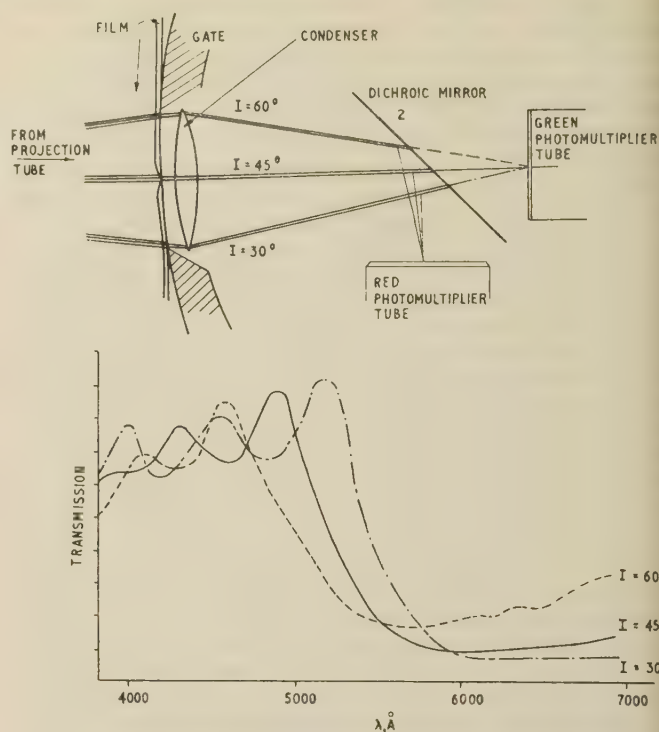
Fig. 5.—Variations in transmission characteristic during passage through gate.

minimum colour flicker and optimum colour rendering from side to side of the field. There are also several possible dispositions of dichroic mirrors in relation to individual transducers, and this choice is beset with many practical considerations regarding the spectral characteristics of (a) the projection phosphor, (b) the possible dichroic reflectors, and (c) the available photo-multipliers. The final layout, which is shown in Fig. 1, possesses the merit that it demands dichroic surfaces which are relatively simple to manufacture. The optical path distance from condenser to photo-multiplier is the same in each case, although, in the interest of a compact assembly, that relating to the blue tube is folded in practice by a fully reflecting mirror.

When this project was emerging from its initial stages, dichroic mirrors of suitable characteristic and sufficient uniform area were not readily available. These reflectors have therefore been made in the laboratory, and novel methods were evolved in the techniques of vacuum deposition and measurement. During evaporation, a light source of constant intensity and chosen colour impinges on a test surface and produces a reflected beam whose brightness is measured by observing the output of a photocell. The time of deposition of each layer is controlled by allowing this reflected brightness to build up to a maximum for the high-refractive-index layers, or to decrease to a minimum for the low. Uniform mirrors are produced by rotating the sample during processing, working at an extremely low pressure, and alternating layers of zinc sulphide ($n = 2\cdot3$) with others of magnesium fluoride ($n = 1\cdot36$).

(3) SCANNING SYSTEM

The scanning cycle requires two components of scanning in the Y-direction at the cathode-ray tube. As shown in Fig. 2, the resultant scanning waveform in 50 c/s television systems derives from 50 c/s sawtooth and 25 c/s shift components, while Fig. 3 demonstrates that, in the case of 60 c/s systems, the waveform is a combination of 60 c/s sawtooth and 12 c/s shift.

It will be noticed that errors in the sawtooth waveforms for line and field affect every patch in the same way, but in the case of the shift waveform, the existence of any error is manifest by a loss of interlace when patches are superimposed at the display tube. Shift waveforms must therefore be stable to within half a line spacing, which, for a 625-line system, requires an accuracy of 1 part in 1 200.

The scanning tube is manufactured with extreme precision to avoid errors in patch shape. These would be introduced, for example, by tilt of the front face (1° tilt produces an interlace error of about four lines) or by errors in gun alignment giving rise to trapezium distortion of the raster and irregularities in spot size. The flat-faced tube is used from optical considerations, but the accuracy of deflection of the spot varies inversely with the cosine of the deviating angle (see Appendix 9). Since this non-uniformity would introduce distortion, by stretching the top of one patch and the bottom of another thereby preventing interlace, corrections to the shift component of the scanning waveform are added to compensate. There are also four magnets, appropriately placed to rectify the corners of the raster, and these secure registration of patches in spite of the usual pincushion deformation of the scan.[7] Care is taken to stabilize the e.h.t. voltage applied to the tube, since instability, by altering the overall scanning sensitivity, would affect both sawtooth and shift components of scanning.

(3.1) Scanning Coils and Circuits

Because of the low frequencies encountered, the high-impedance field-scanning coils are supplied by direct valve drive instead of transformer coupling[8] (Fig. 6). The required waveform
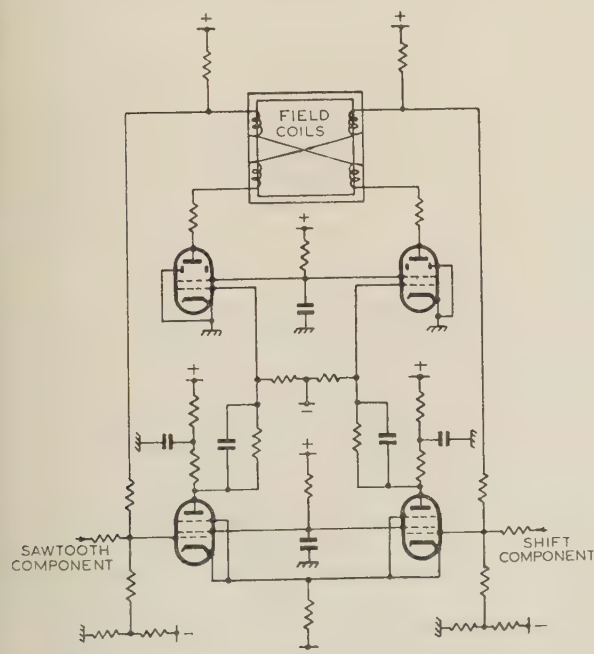
Fig. 6.—Field drive circuit.

is obtained by generating separately the sawtooth and the shift waveform, the two being added in a mixer circuit which drives the push-pull output stage. The push-pull circuit avoids problems introduced by direct current through the coils, since these are arranged so that the magnetic fields due to the standing current of each valve in each limb of the coil assembly are cancelled. Negative feedback, controlled by the current through the

scanning coils, is introduced to limit valve distortion and to maintain linearity with stability.

Line and field scanning coils are wound on a square yoke of Ferroxcube I sections, each coil being provided with a separate micrometer movement in a transverse direction; this allows adjustment for the best patch shape under working conditions (Fig. 7). The assembly is located about an axial Paxolin tube
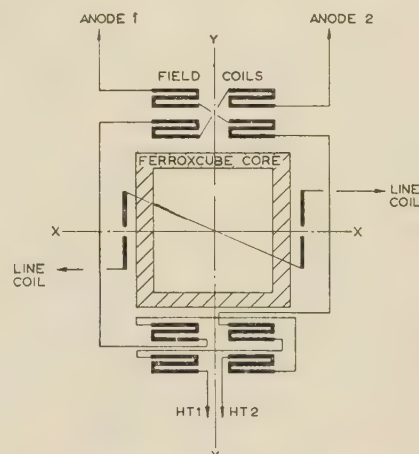
Fig. 7.—Scanning coils. Arrangement XY.

by accurately machined spiders to which the Ferroxcube pieces are pressed, and then the yoke is enclosed in a thin copper shield to increase efficiency at line frequency. The focus coil has a soft-iron shroud and requires an m.m.f. of 800 AT, and, to correct for possible gun misalignment, there are centring coils at the rear of the assembly.

It is automatically ensured that, when the machines start, the patch motion is synchronized with a film frame. Otherwise, in a 50 c/s system, the frame bar could occupy one of two positions depending upon the particular moment of switching on. In this case, uncertainty is avoided by generating a phasing pulse at the motor-drive unit every second film frame (four television fields), and using it to trigger the patch cycle. The square wave is controlled in amplitude by passing through a two-diode window, and connected to the mixer circuit to bring the patches into interlace.

When dealing with the 60 c/s system, which could give rise to five possible positions of the frame bar depending upon the precise moment of switching, an interlocked 12 c/s shift waveform is required. This is provided by relays which follow a regular sequence and select potentials from a high-stability resistance bleed from the h.t. line to earth (Fig. 8). The step voltages, which are adjustable over a range of $\pm 5\%$ by series potentiometers, are connected via series resistors to the contacts of polarized high-speed relays which have two windings and require to be driven both on and off. The relays are operated by switching pulses generated by a cascade of five multivibrators. The phased pulse from the motor-drive equipment is applied to a 1-in-5 counter circuit, which divides the field repetition rate of 60 c/s and produces a 12 c/s output to trigger the first multivibrator. The pulse then generated is used to switch the respective relay for one field period; its back edge is used to trigger the second multivibrator, which also generates a pulse of one field period, and so on in immediate succession.

(3.2) Automatic Control of Scanning-Tube Brightness

The brightness of a scanning tube with fixed beam current varies over the raster because the spot size, and therefore current
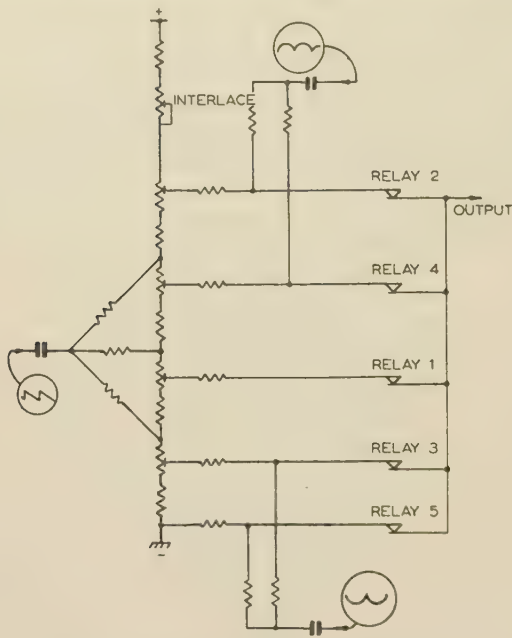
Fig. 8.—60 c/s shift waveform generator.

density at the phosphor screen, is not constant. Unfortunately, zinc oxide exhibits a saturation effect which limits brightness as the current density increases, so that, in areas of smallest spot size, the intensity is less. Brightness is also affected by browning of the front glass plate, owing to electron penetration, and this causes a loss of transmission which may not be uniform for different patch positions. If succeeding picture fields have different brightnesses flicker is apparent, and so a photo-multiplier tube is arranged to measure the brightness at every point of the raster, and pass an error signal to the grid of the scanning tube (Fig. 9). Because slow changes are experienced, a bandwidth of

Fig. 9.—Automatic control of scanning-tube brightness.

150 kc/s, which is ten times the highest line frequency, is sufficient, and the photo-multiplier output signal, limited to 1 mA, feeds a 150-ohm low-capacitance cable with a first-order doublet network termination. At frequencies within the 150 kc/s band, the impedance is 1·5 kilohms, giving 1·5 volts signal, but at high frequencies the line is terminated in its characteristic impedance of 150 ohms. It should be noted that variations in the peak signal only are wanted, and the flyback pulse is 'backed off' by an

opposite pulse. A 3-stage wide-band amplifier provides sufficient voltage swing to reduce brightness variations by 15 dB.

(4) FILM TRACTION

In order to fulfil the scanning requirements, the film speed as it traverses the scanning zone must be maintained to within very close limits. For example, to secure vertical stability of a quarter line with a 625-line system, errors in the instantaneous position for 35 mm film must not exceed 0·000 3 in, and in the case of 16 mm film, 0·000 12 in. These demands are particularly severe in the case of the 16 mm film, which is located by only one perforation per frame, and yet is submitted to magnification 2·15 times greater than the 35 mm film, which has eight perforations available. In each case the speed of transport must be uniform to a very high degree, and consequently the machines require to be built with considerable precision.

In any mechanism, accuracy of performance is generally inversely proportional to the degree of complexity, and for this reason the gearing in this machine is limited to one stage per function (Fig. 10). The speed reduction between driving motor and film sprockets is considerable, and so a satisfactory choice of gearing methods is to use the worm and wormwheel, which can be made in hardened form with great accuracy by modern techniques of thread and profile grinding. Further refinement in performance is obtained by providing, between the wormwheel and the master-sprocket spindle, a mechanical filter system which is the analogue of an electric low-pass filter.

The mechanical filter consists of a Duralumin drum, inside which is concentrically mounted a freely-rotating, heavy flywheel. Since only small clearances are allowed between the drum and flywheel and the space is filled with oil of a suitable viscosity, any rotational oscillation between the two tends to be damped. The drum itself is solidly fixed to the sprocket spindle; it is driven by four cantilever springs, attached to the wormwheel boss, which bear against adjustable abutments on the drum. The chief requirements for correct operation of such an arrangement are, first, the elimination of all backlash or play between worm and wormwheel operating surfaces, and secondly, the provision of a reasonably constant load on the master film sprocket. These features can be regarded respectively as the earth and terminating resistance of the filter system.

Backlash is avoided by making the wormwheel boss relatively large in diameter and arranging for it to rotate upon a stationary quill surrounding the bearings of the master-sprocket spindle. The cylindrical surface of this quill is lapped to suit the bore of the wormwheel and to provide a small clearance. Consequently, when the wheel rotates, the resistive load presented by shearing the oil film between these surfaces effectively maintains contact between the driving faces. The independent support provided by the quill also relieves the sprocket spindle from any radial loading due to worm and wheel reactions. The second requirement of load on the sprocket is applied by friction of the film over the curved film gate, the pressure between film and gate being determined by the effect of a regulating roller above the gate. The enlarged spindle of this roller rotates in a bore with small clearance, and, since this is oil-filled, a resistive load is imposed by shearing the oil film in the same way as for the master-sprocket arrangement.

The master film sprocket is of special construction, incorporating hardened and profile-ground toothed rings to engage the film perforations. Their design allows either new or shrunk film to be run without sprocket-tooth interference. The usual feed and pull-off sprockets are provided, the film being kept taut between feed sprocket and the regulating roller by means of a spring-loaded jockey roller.

FROM FILM SPOOL

PULL-OFF SPROCKET

FEED SPROCKET

FILM-TENSION ROLLER

PHOTO-MULTIPLIER TUBES

OIL-FLOW INDICATOR

LOOP TENSION ROLLER

VISION-FILTER FLYWHEEL

COLOUR ANALYSER

FILM GATE

LENS ASSEMBLY

VISION MASTER SPROCKET

SOUND-TRACK OPTICS

SOUND-FILTER FLYWHEEL

BELT TO FILM TAKE-UP SPOOL

FILM DETOUR ROLLER

SOUND-EXCITER LAMP HOUSE

SOUND DRUM

SOUND PULL-OFF SPROCKET

FILM HOLD-BACK SPROCKET

FILM FRAMING KNOB

OIL-LEVEL INDICATOR

TO FILM SPOOL

SELSYN MOTOR

Fig. 10.—Mechanism of 35 mm scanner.

The 16 and 35 mm machines are very similar except for the size of sprockets and layout of spindles, the same basic castings and many other details being common to both types. Each is driven from a Selsyn motor mounted close beneath the mechanism and coupled directly to the main shaft. To align the film in the gate for 'framing', the motor body is given a limited rotational movement around its axis, this adjustment being effected by means of a worm engaging with a toothed segment on the motor case.

(4.1) Motor Selsyn Set

The individual Selsyn motors are supplied from separate generating units housed in cubicles. These comprise a 3-phase motor of special type driving a main Selsyn generator through suitable gearing; the output is then available to the motors of

film mechanisms and also to associated sound equipments which are required to operate in synchronism. Although an optical sound head is provided on each picture unit, it is sometimes convenient to use an independent sound film or magnetic tape, and by this method a number of picture and sound reproducers can be operated in parallel.

Each auxiliary driving motor must be synchronized with television field pulses and independent of the mains supply frequency; this allows film to be inserted into programme derived from another source, even though there may be a frequency difference of up to ± 2 c/s.

The main generator driving motor has 3-phase windings on both stator and rotor, the stator winding being connected directly to the station mains supply, while the rotor is fed from a separate source controlled by the television field pulses. The speed ratio

between driving motor and Selsyn generator can be arranged to provide the correct film speed at supply frequencies of either 50 or 60 c/s, by using an appropriate set of gears. A contact-maker in this drive provides a timing pulse to maintain synchronism of the patch movement with film motion, as explained in Section 3.1.

(5) COLOUR ANALYSIS

(5.1) Choice of Co-ordinates

The following considerations are important when adopting basic colour co-ordinates for a system:

(a) Fidelity of colour reproduction.
(b) Signal/noise ratio in the luminance signal.
(c) Efficient use of light.

The discussion will be limited to two possible systems of analysing co-ordinates (see Fig. 11), namely:

(i) The N.T.S.C. primary colours, which are RGB (0·67, 0·33) (0·21, 0·71) (0·14, 0·08).
(ii) The C.I.E. primary colours, which are XYZ (1, 0) (0, 1) (0, 0).
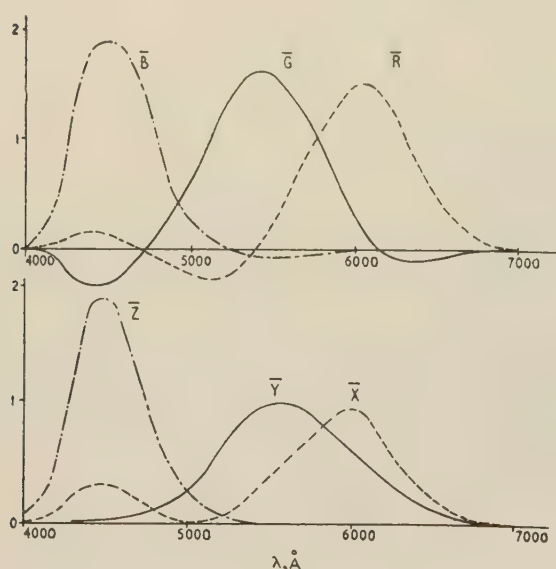
Fig. 11.—RGB and XYZ distribution coefficients.

(5.1.1) Colour Fidelity.

Dealing first with the RGB co-ordinates, it will be seen that there are negative components in the required responses $\bar{R}$, $\bar{G}$, $\bar{B}$, and so it is necessary to employ at least six transducers, i.e. two or more photo-multipliers per colour lane, to synthesize these curves accurately. If three photo-multipliers only are used, the colour analysis becomes inaccurate, because the minor negative lobes of the curves are neglected. As an example, assuming otherwise complete analysis of some representative Wratten filters, errors which would be introduced have been computed and plotted in Fig. 12. It will be seen that, generally, colours are desaturated, and some examples exhibit an error about six times greater than that which is accepted as photometrically just perceptible.

An alternative is to use XYZ co-ordinates, since the curves are entirely positive, and a complete and faithful analysis of the object colours can be achieved with three transducers. In this case, it would be necessary to obtain RGB signals before gamma correction by including matrix circuits.
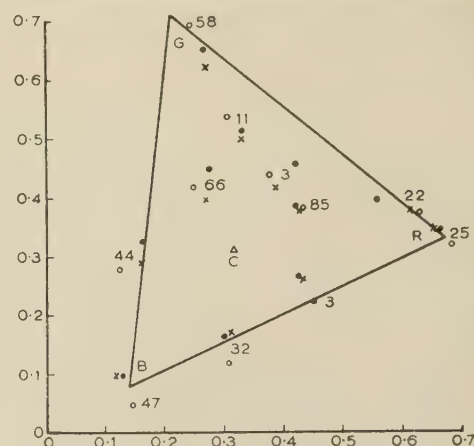
Fig. 12.—Errors in RGB analysis of specimen filters.

● Measured in channel.
× Calculated (neglecting negative parts).
○ Actual value for colour filter used.
△ Illuminant C.

(5.1.2) Signal/Noise Ratio.

The luminance, or output-brightness, signal is compounded to the following specification:

$$V'_Y = 0\cdot3V'_R + 0\cdot59V'_G + 0\cdot11V'_B$$

and it is important to investigate how the r.m.s. noise voltage of each component combines to produce noise in the luminance signal.

In a flying-spot scanner, the noise produced is almost entirely due to random fluctuations in electron emission from the photo-cathodes, and so is proportional to the square root of the signal current. However, the gamma-correction circuit, to which the output signal is applied, also has an approximately square-root relationship, and the combination produces a constant noise voltage in each colour lane. This is true for all signals except those of low intensity, where the gamma-correction circuits depart from theoretical performance. However, since

$$V'_{NY} = \sqrt{[(0\cdot3V'_{NR})^2 + (0\cdot59V'_{NG})^2 + (0\cdot11V'_{NB})^2]}$$

where V'_{NY}, V'_{NR}, etc., represent noise voltages after gamma correction in V'_Y, V'_R, etc., signals, respectively, it will be seen that the noise voltage, V'_{NY}, in the luminance signals is independent of colour. This means that, in the RGB system, the r.m.s. noise voltage remains nearly the same for white as it is for colours, apart from the most saturated ones, where noise is reduced.

In the XYZ system, matrices are required to obtain RGB signals before gamma correction can be applied. For white, the signals voltages are applied to the same point of the gamma curve in each colour lane, so that noise voltages are modified in the same way, and the V'_Y signal reverts to the signal and noise voltages from the Y photo-multiplier tube. However, for any *colour*, the derived RGB signals are applied to different points of the gamma curve in each colour lane, and the noise voltages are modified in different ways. Therefore, the output V'_Y signal for any colour contains components of X and Z signals, and the V'_Y noise voltage contains noise from the X and Z multiplier tubes. The result is that a red field, for example, may have as much as 20dB more noise than a white one. The r.m.s. noise voltages for different colours in both systems are shown in Figs. 13(a) and 13(b). These computations were derived from the following values, which approximate to those realized in a practical colour scanner:

X and R noise signals = 50mV in a 1-volt signal.
Y and G noise signals = 25mV in a 1-volt signal.
Z and B noise signals = 50mV in a 1-volt signal.

(a)

(b)
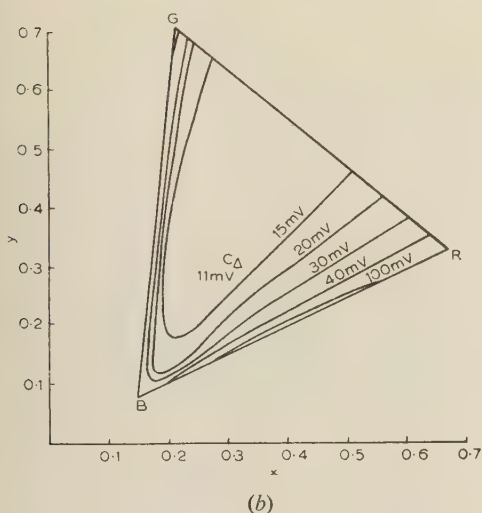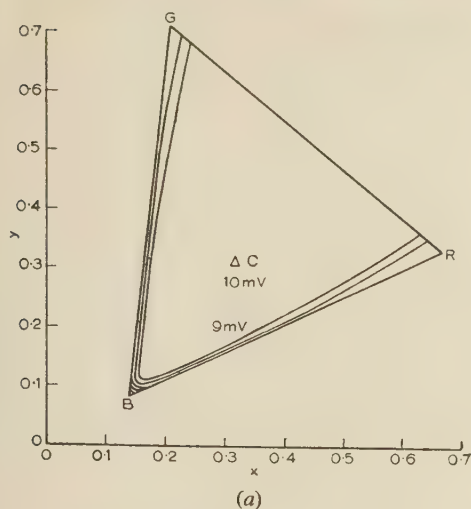
Fig. 13.—R.M.S. noise (in millivolts) for 1 volt (peak) V'_Y signal.

△ Illuminant C.

(a) Using RGB co-ordinates. R.M.S. noise in RGB signals is
$R = B = 50$ mV.
$G = 25$ mV.

(b) Using XYZ co-ordinates. R.M.S. noise in XYZ signals before matrix to RB is
Z and $X = 50$ mV.
$Y = 25$ mV.

(5.1.3) Utilization of Light.

The phosphor of an ideal scanning tube would produce light energy throughout the entire visible spectrum, but would also exhibit a short decay time—in practice, less than 1×10^{-6} sec. Unless this time is realized, the light emitting from a point in the phosphor after the electron beam has passed degrades picture quality owing to continuous motion of the film. The most satisfactory screen material is zinc oxide (approximately equivalent to P.15), and this is used in the scanning tube type R5161U. From the available photo-multiplier tubes, two types have been chosen for their particular spectral characteristics. The first type 6097, which uses an antimony-caesium photo-layer sensitive to blue and green but offering no response beyond 6000Å, and the second is type 6095 using a silver-bismuth layer, which extends its response to 7000Å, although it is less sensitive in the blue and green region. The spectral characteristics of the phosphor and of these two photo-layer materials is shown in Fig. 14.
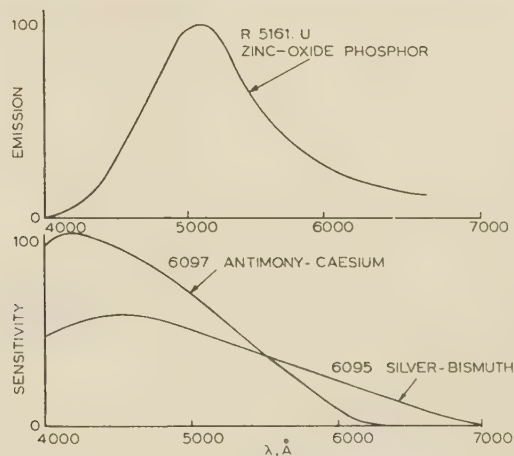
Fig. 14.—Spectral characteristics of phosphor and photo-layers.

Reference to Fig. 11 shows that $\bar{R}$, $\bar{G}$, and $\bar{B}$ cover narrower spectral bandwidths and overlap less than $\bar{X}$, $\bar{Y}$, and $\bar{Z}$. It is evident, too, that the antimony-caesium layer used for $\bar{B}$ and $\bar{Z}$, respectively, also exhibits a characteristic closer to $\bar{G}$ than the silver-bismuth layer which is required for $\bar{Y}$. The latter type is, however, eminently suitable for $\bar{R}$ and $\bar{X}$, which are similar in shape. Since the use of a photo-multiplier with more sensitivity endows the G signal with at least 3dB less noise than the Y signal, and since the colour discrimination is improved, the use of RGB co-ordinates is considerably more efficient.

(5.1.4) The Choice Between RGB and XYZ Systems.

Colour analysis, which has perfect fidelity in an XYZ system, suffers noticeable distortion with the RGB system if three photo-multipliers only are used. However, it has been shown on the baisis of signal/noise ratio in the luminance signal, that the adoption of RGB colour co-ordinates is preferable to XYZ. A further factor concerns the types of photo-multiplier used; the combination which reduces noise in the G signal offers to the RGB system an increase in light efficiency which could not otherwise be achieved. Additional points in favour of this system are the convenience of amplifier design and operation, and the fact that matrices are avoided.

Therefore, because of the foregoing factors, the RGB system of taking co-ordinates was chosen and the rest of the paper deals with problems consequent upon this decision.

(5.2) Beam Splitter

An efficient beam splitter would divide light at every wavelength in accordance with the spectral requirements of each colour lane. If the light outputs to each colour lane are $\bar{r}V_\lambda$, $\bar{g}V_\lambda$ and $\bar{b}V_\lambda$,

$$\bar{r}V_\lambda = \frac{\bar{R}}{\bar{R} + \bar{G} + \bar{B}}V_\lambda$$

$$\bar{g}V_\lambda = \frac{\bar{G}}{\bar{R} + \bar{G} + \bar{B}}V_\lambda$$

$$\bar{b}V_\lambda = \frac{\bar{B}}{\bar{R} + \bar{G} + \bar{B}}V_\lambda$$

which assumes that $\bar{r} + \bar{g} + \bar{b} = 1$ and that the R, G, and B channels are of equal importance.

However, the N.T.S.C. system specifies the luminance signal, so that, supposing light to be used with equal efficiency in each colour lane, the relative distribution of light at the light divider

should be $0 \cdot 3R$, $0 \cdot 59G$, $0 \cdot 11B$. The output characteristics of the beam splitter are therefore

$$\bar{r}V_\lambda = \frac{0 \cdot 3\bar{R}}{0 \cdot 3\bar{R} + 0 \cdot 59\bar{G} + 0 \cdot 11\bar{B}}V_\lambda$$

$$\bar{g}V_\lambda = \frac{0 \cdot 59\bar{G}}{0 \cdot 3\bar{R} + 0 \cdot 59\bar{G} + 0 \cdot 11\bar{B}}V_\lambda$$

$$\bar{b}V_\lambda = \frac{0 \cdot 11\bar{B}}{0 \cdot 3\bar{R} + 0 \cdot 59\bar{G} + 0 \cdot 11\bar{B}}V_\lambda$$

Reference to Fig. 1 will show that the chosen arrangement of dichroic mirrors requires the following relationship to be fulfilled:

If the reflection from dichroic mirror 1, R_1, is $\bar{b}$, the transmission of the same mirror, T_1, is $(1 - \bar{b})$

and

$$T_1 R_2 = \bar{r}$$

so that

$$R_2 = \frac{\bar{r}}{1 - \bar{b}}$$

and

$$T_2 = 1 - \frac{\bar{r}}{1 - \bar{b}} = \frac{\bar{g}}{1 - \bar{b}}$$

Transmission curves for mirrors 1 and 2, comparing the ideal requirements with the practical results achieved, are shown in Fig. 15.
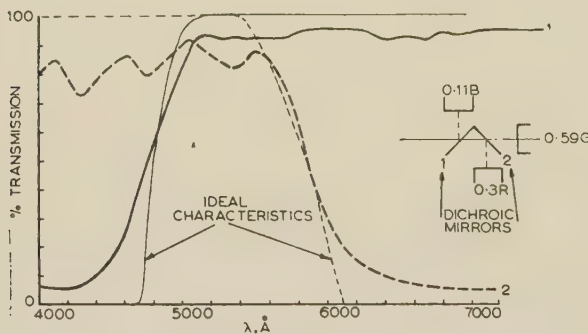
Fig. 15.—Beam-splitter characteristics.

Dichroic mirror 1.
——— Ideal curve.
▬▬▬ Curve obtained.
Dichroic mirror 2.
– – – Ideal curve.
▬ ▬ ▬ Curve obtained.

(5.2.1) Colour Correcting Filters.

The products of the spectral characteristics of all the elements in the colour lanes should produce the required distribution coefficients $\bar{R}$, $\bar{G}$ and $\bar{B}$, respectively (Fig. 16).

$$\bar{R} = \bar{F}_R \times \bar{S}_R$$
$$\bar{G} = \bar{F}_G \times \bar{S}_G$$

and

$$\bar{B} = \bar{F}_B \times \bar{S}_B$$

where $\bar{S}_R$, $\bar{S}_G$ and $\bar{S}_B$ represent the products of the spectral characteristics of all the elements,

and $\bar{F}_R$, $\bar{F}_G$ and $\bar{F}_B$ are correcting filters required in the R, G and B lanes, respectively.

For example, as applied to the red lane, $\bar{S}_R$ is the product of spectral characteristics of (*a*) the scanning tube, (*b*) the transmission of mirror 1, (*c*) the reflection from mirror 2, and (*d*) the red photo-multiplier tube. In a similar manner the products $\bar{S}_R$, $\bar{S}_G$ and $\bar{S}_B$ can be computed, allowing, where necessary, for
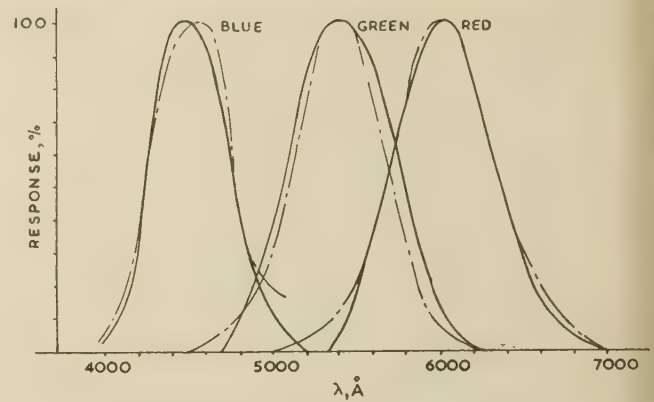
Fig. 16.—Spectral characteristics of RGB lanes.
—·— Curves obtained.
——— Ideal curves (RGB without negative lobes).

the small additional loss in the blue region due to absorption in the glass.

The optical system of the flying-spot film scanner is designed to produce a diffuse beam of light at the photo-multiplier surfaces, and a novel design for the practical realization of correcting filters has been evolved. These are constructed with calculated amounts of commercially obtainable gelatine colour filter from which the required transmission can be accurately and most efficiently synthesized. By inspection of the required overall characteristic, in comparison with the products of $\bar{S}$ and probable individual gelatine filters, appropriate sectors may be arranged about the centre of the light beam, offering colour analysis independent of the aperture of the system. A simple correcting filter for the 'red' lane, demonstrating the method by which it has been constructed, is shown in Fig. 17.
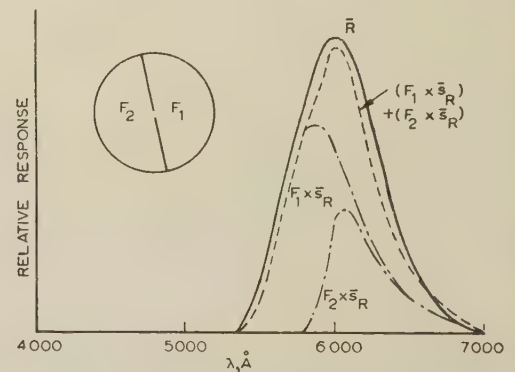
Fig. 17.—A correcting filter for 'red' lane.
F_1 and F_2 are filters.

The overall response curves for the three colour lanes demonstrate that the results are a compromise between accuracy of analysis and efficiency of light transmission (Fig. 16). For example, the 'green' curve could be improved at the red end only by increasing the amount of red filter, but this would reduce the light transmission and increase noise in the 'green' signal.

(5.2.2) Test Procedures.

There are two convenient ways of assessing fidelity of colour analysis. In the first method the response of each colour lane to a number of narrow-band interference filters is measured. The results obtained in this way are presented in Fig. 16. The

ᴝond method derives the chromaticity co-ordinates of a par-
ᴉᴉlar gelatine filter from measurements of the RGB signals.
ᴉthis way the 'measured' values were obtained for the Wratten
ᴇᴇs shown in Fig. 12.

(6) VIDEO-FREQUENCY SIGNALS

(6.1) RGB Signals

Each projector has its own colour picture head, from which
ᴇ output signals are approximately proportional to the incident
ᴉours.   The outputs from one colour picture source are selected
ᴉ relays and passed to a vision-level control, which is an
ᴄurately matched 3-ganged potentiometer, maintaining approxi-
ᴉtely 75 ohms impedance over a signal variation of 2 : 1 in
ᴄh colour lane.   The RGB signals at this point are 18mV
ᴇak) colour while black level is at zero volts (Fig. 18).   Correc-

Fig. 18.—Video-frequency block schematic.

ᴘns for afterglow of the scanning tube are made before
ᴘnnection to the amplifier unit, which delivers signals at a level
ᴉitable for gamma correction.   A variable lift pulse is added
ᴇ each colour signal in this amplifier unit, so that picture black
ᴉay be moved towards black level.   Gamma correction is applied
ᴘarately to each colour signal, and it can be varied to adjust
ᴇ contrast range of the RGB colour signals to suit the colour
ᴉᴘlay.   Suppression pulses are then added to provide RGB
ᴉtᴘut signals suitable for monitoring, white balancing, etc.   A
ᴘᴉᴏur-mask unit may be inserted at this point in the vision
ᴉ꜀uit, to derive a new set of colour co-ordinates R'', G'', B''
ᴉᴇre:

$$R'' = R' + b_1(G' - R') + c_1(B' - R')$$
$$G'' = G' + a_2(R' - G') + c_2(B' - G')$$
$$B'' = B' + a_3(R' - B') + b_3(G' - B')$$

The coefficients b_1, c_1, a_2, c_2, a_3, and b_3 are variable, but when
reduced to zero, they still offer an unimpaired relationship for
white:

$$R'' = R', G'' = G', \text{ and } B'' = B'$$

The colour-mask unit offers control of hue and saturation of
the reproduced colours to provide a more pleasing effect, or to
correct for shortcomings in the rendering of some types of colour
film.   For example, a more saturated red is obtained by
increasing the coefficients a_2 and a_3, thus reducing the 'green'
and 'blue' components.   On the other hand, increasing a_2 and
reducing a_3 will change the hue of colours with a red content,
because the 'green' signal is reduced but the 'blue' is increased.

(6.2) Colour Modulator

The colour modulator accepts the output RGB signals and
produces a composite signal to the relevant specification, which
may be for the 525-line 60-field/sec, 405-line 50-field/sec or
625-line 50-field/sec television systems, respectively.   Variable
aperture correction is applied to the luminance channel (Y) only;
this compensates for loss of high-frequency response at the
scanning tube and optical components, and a maximum of 9 dB
boost at 5 Mc/s is available.   Synchronizing pulses are added to
the Y signal, and the bandwidths of the two colour difference
signals, I and Q, are appropriately limited.   One set of YIQ
signals is then available for connection to a local colour monitor,
and from a second set with an externally generated colour sub-
carrier, the composite colour signal is produced.

(6.3) White Balance

The N.T.S.C. system requires that the chrominance signal
disappear, i.e. $R = G = B$ for Illuminant C (6500°K), although,
when clear film base is projected in an average cinema, the colour
temperature of white reflected from the screen is 5200°K.   In
the colour scanner it is simpler to set up white balance if equal-
energy white is assumed, i.e. 5500°K, in which case, RGB
video-frequency signals should be adjusted with clear film base
to give

$$R = 1 \cdot 09 \text{ units of signal.}$$
$$G = 0 \cdot 99 \text{ units of signal.}$$
$$B = 0 \cdot 84 \text{ units of signal.}$$

(6.4) Monitoring Facilities

Complete video monitoring facilities are included, a high-grade
monochrome picture display being provided on a 14 in rect-
angular tube, with a separate waveform display presented on a
4 in tube.   By pressing the appropriate button at the control
panel it is possible to view any one of the R, G, B or Y signals.
In this way white balance, and corrections for afterglow or tilt,
can be observed separately while being applied to the individual
colour lanes.

(7) ACKNOWLEDGMENTS

(5) REFERENCES

(1) CONDLIFFE, G. E.: British Patent No. 475032, April, 1936.
(2) HOLMAN, H. E., and LUCAS, W. P.: 'A Continuous-Motion System for Televising Motion-Picture Film', *Proceedings I.E.E.*, Paper No. 1316 R, April, 1952 (**99**, Part IIIA, p. 95).
(3) NUTTALL, T. C.: 'The Development of a High-Quality 35mm Film Scanner', *ibid.*, Paper No. 1275 R, April, 1952 (**99**, Part IIIA, p. 136).
(4) JESTY, L. C.: British Patent No. 608939, October, 1944.
(5) GOLDMARK, P.: U.S. Patent No. 261848.
(6) QUINN, S. F., and PYKE, P. J.: British Patent Application No. 32723/55.
(7) BULL, E. W., and PEMBERTON, M. E.: British Patent No. 605572, July, 1945.
(8) HADFIELD, L. D.: British Patent No. 705197, June, 1951.
(9) BINGLEY, F. J.: 'Colorimetry in Colour Television—Part III', *Proceedings of the Institute of Radio Engineers*, 1954, **42**, p. 51.
(10) WRIGHT, W. D.: 'The Measurement of Colour' (Adam Hilger, London, 1944).
(11) BREWER, W. L., LADD, J. H., and PINNEY, J. E.: 'Proposed Controls for Electronic Masking', Institute of Radio Engineers Convention Record, 1955, Part 7.
(12) DENISON, R. C.: 'Aperture Compensation for Television Cameras', *R.C.A. Review*, 1953, **14**, p. 569.
(13) CORNWELL CLYNE: 'Colour Cinematography' (Chapman and Hall, 1951).
(14) QUINN, S. F.: British Patent Application No. 32724/55.

(9) APPENDIX

Deflection Sensitivity Error due to a Flat Tube Face

It is assumed that the scanning coils are ideal, i.e. the scanning field is uniform and directly proportional to the scanning current I.

In Fig. 19 the beam will follow a circular path of radius r while in the field H, where

$$r = \frac{mv}{eH}$$

m, e and v being the mass, charge and velocity of the electron, respectively. Also H is proportional to I as stated.

i.e.
$$r = \frac{K}{I} \quad \ldots \quad \ldots \quad \ldots \quad (1)$$

where K is a constant

From Fig. 19
$$\left. \begin{array}{l} D = L \tan \theta \\ r = l/\sin \theta \end{array} \right\} \quad \ldots \quad \ldots \quad (2)$$
and

The deflection sensitivity is given by

$$D/I = \frac{Lr \tan \theta}{K}$$

$$= \frac{Ll}{K} \frac{1}{\cos \theta} \quad \ldots \quad \ldots \quad (3)$$

after substitution from eqn. (2).

From eqn. (3) the deflection sensitivity at the centre of the
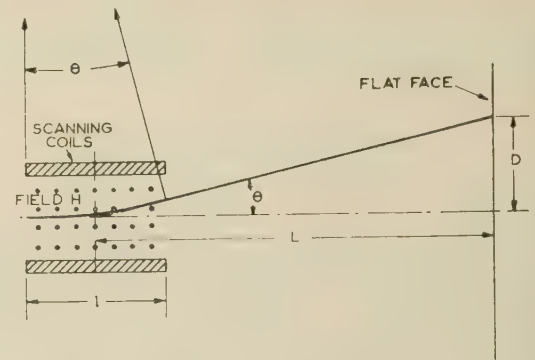
Fig. 19.—Deflection sensitivity errors due to a flat-faced tube.
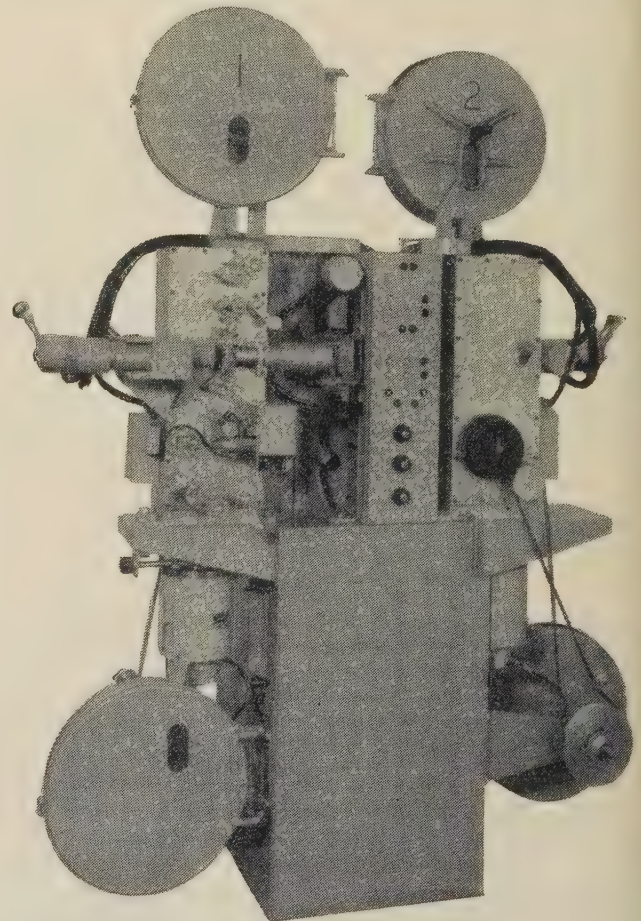
Fig. 20.—Scanner assembly.

tube face is Ll/K. Eqn. (3) shows that the deflection sensitivity is not constant but varies as the inverse of the cosine of the deflection angle.

The percentage error in the deflection sensitivity, for a deflection θ, may thus be given approximately by $(1 - \cos \theta)100$.

Mr. G. G. Gouriet: The continuous-motion approach to film-scanning, in which the image is immobilized by the motion of the scanning spot, is still, in my opinion, the most elegant of all methods, and it is pleasing to see this method adapted so admirably to the problem of film scanning in colour television. It is clear that the authors have encountered a number of interesting and difficult problems, the solution of which has required a considerable measure of ingenuity.

I am particularly interested in Section 5, in which the authors consider the possibility of obtaining faithful colour analysis by analysing in terms of the C.I.E. co-ordinates for which no negative responses are required. They arrive at the conclusion that the XYZ axes are inferior to RGB axes in respect of the signal/noise ratio, and this is attributed to instrumental causes associated with the need for gamma correction after matrixing.

I find it difficult to understand why the choice of analysis co-ordinates or the gamma-correction circuits should have an effect on the signal/noise ratio of the reproduced picture. A change of axis involves only a linear transformation, whilst any effect produced by gamma correction is finally removed by the inverse gamma of the display tube.

I suggest that the decrease in signal/noise ratio which results from using the XYZ co-ordinates can be explained as follows. Perfect colour analysis in terms of the RGB primary signals in question requires regions of negative response in the analysis characteristics. Since 'negative' light has no physical meaning, any practical method of achieving perfect colour analysis necessarily involves analysis in terms of positive quantities, with subsequent subtraction of electrical signals in order to obtain the desired responses. Since the subtraction of signals cannot be performed before noise has been added, the result of such subtraction is always a decrease in signal but an increase in noise. This follows from the fact that the noise power of incoherent sources always adds irrespective of whether the source outputs are added or subtracted. It would seem, therefore, that the decrease in signal/noise ratio is the price which must inevitably be paid for improving on the standard of colour reproduction which is obtained when the negative regions of the analysis curves are neglected. In other words, I suggest that this is an example of the general principle by which information can be exchanged for signal/noise ratio.

I believe that the result would be the same even if we had 'magic' filters with negative transmission (whatever that may mean) in the appropriate regions; the noise power associated with these regions would still have to be 'added', and so the signal/noise ratio would be worse than if the negative regions had been neglected.

Dr. A. R. A. Rendall: The authors seem to have entered into the arena of controversy, as the equipment is designed for 525-line working.

There seem to be two broad schools of thought—the use of three vidicon cameras or the flying-spot arrangement. The protagonists of the vidicon cameras claim that better signal/noise ratio is obtained, whereas the protagonists of the flying-spot arrangement claim that it overcomes some very difficult registration problems. It would be interesting to know the impact of these proposals in the United States, because I believe that the three-vidicon-camera system is still favoured there.

I am particularly impressed by the problem stated in Section 2, i.e. the need for a precise geometrical shape for the scanning patch, irrespective of its position on the scanning tube. The principles and the method employed require that there shall be scanning patches for odd and even fields located at different parts of the tube. The familiar pincushion effect must there-fore be avoided at all costs and to a very high degree of precision.

It is stated in Section 3 that the shift waveforms must be stable to within half a line spacing. I should have thought that the whole scanning-mechanism shift-waveform geometrical accuracy must be such as to produce the odd and even film-frame scans to within this degree of accuracy.

The solution of the problem of the feedback and screen brightness given in Section 3 is very elegant, but it seems rather complex. Will the required accuracy be maintained in service, and, if not, will there be flicker? Is there any loss of performance in this respect if the scanning tubes are changed during service?

The paper presents an overall picture of a very exacting problem which has been tackled with great skill and infinite attention to detail, but I wonder whether all this meticulous accuracy can be maintained in regular everyday use. Although it is very difficult in such a paper to convey an idea of the overall performance, I wonder whether it is possible to express in any satisfactory terms the overall accuracy of registration in the various parts of the picture. It is very difficult to get any mental picture of what is happening from the square diagram. For example, with regard to signal/noise ratio, it is difficult to appreciate the net result until we see the picture itself.

Mr. L. C. Jesty: In Section 5.1, which deals with accuracy of colorimetry, there is still something which eludes me. This may be because accuracy of colorimetry is not vital, as is well known in the photographic industry. Picture quality is much more vulnerable to gamma variation, and in television apparently to noise. However, I recall a demonstration of the authors' film scanner last year using XYZ analysing primary signals, in which the picture was noise-free when the XYZ signals were routed separately. It was only when they were coded into an N.T.S.C.-type signal that the noise showed up on a saturated colour. If the upper luminance frequencies are to be used in the N.T.S.C. system, it appears to be vulnerable to crosstalk noise between luminance and chrominance. Whilst the choice of RGB analysing primary signals may be a better engineering solution in the authors' case, this may not always be so.

So long as RGB signals are ultimately required by the receiver display, then, for minimum noise from the pick-up device, the best choice of analysing primary responses would appear to be those requiring minimum dilution by subsequent matrixing. A set of real analysing primary responses (i.e. no negative values and therefore capable of accurate colour reproduction) which may have advantages can be derived from Fig. B of my contribution to a previous discussion.* This shows a projection of the C.I.E. spectral locus on to a constant-luminosity plane with a triangle fitting very closely to the locus. The characteristics of the analysing primary responses for this close-fitting triangle can be derived from the co-ordinates of the corners and the spectral locus in the usual way. Limiting the reproduced brightness range to 100 : 1 reduces the area of the diagram required very considerably, as shown by the shaded area in Fig. B. This should enable the analysing primary responses to be brought still closer to the reproducing primary responses.

With regard to the method used in this film scanner for immobilizing the film image, I have always felt that the requirement for accurate superimposition of the successive positions of the scan on the flying-spot tube, coupled with accurate matching of brightness all over the picture area, was a formidable task. However, I have seen some very excellent pictures from this machine which justify the method. Are there any serious

* *Proceedings I.E.E.*, 1949, **96**, Part II, p. 467. The footnote in the Reference should be SMITH, T.: Report of Joint Discussion on Vision by the Physical Society and Optical Society, 3rd June, 1932, p. 218.

difficulties with tube replacement, tube life, screen ageing and similar factors relating to the continued generation of a good signal after the initial setting-up has been completed?

Mr. W. P. Vinten: I am interested in the possible unsteadiness caused by the relatively wide tolerances which are necessarily allowed in spacing of film perforations. This applies particularly to the 16mm gauge, in which, owing to the high magnification, errors in registration quickly become objectionable. To appreciate how these errors can adversely affect the steadiness of the flying-spot machine, it is necessary to understand the process of film making.

The raw stock consists of a length of plastic material perforated normally on one side only. Largely owing to the plastic nature of the material, these perforations are almost certainly not exactly in their correct position. When this film is exposed in a cinematograph camera it is held in register, possibly by means of a register pin. The hole used for registration purposes will probably be one frame away from the exposing aperture. If the camera is in perfect condition it places the exposures on the film in exactly the right position relative to the hole used for registration purposes. Therefore the exposures have the same errors as are present in the perforations. If care is taken to ensure that these same holes are used for registration throughout the intermediate printing processes and also for final projection, all errors cancel and the result is a perfectly steady picture.

However, the flying-spot machine does not use any specific hole for registration purposes, but is a constant-velocity machine and ensures that the film, when passing the gate, has a perfectly uniform velocity.

As the exposures are not accurately spaced, it is probable that unsteadiness will occur. The ideal conditions on a flying-spot machine are that exposures should pass the aperture at intervals of 1/25 sec. This cannot be the case with inaccurately spaced exposures, however constant the velocity of the film.

It would be interesting to know whether it is possible to overcome this problem or if the defect is so small as to be unnoticeable.

Mr. T. Jacobs: It is of paramount importance that a flying-spot scanner produce video-frequency signals with a good signal/noise ratio. The somewhat inefficient use made of the light from the flying spot is therefore a little surprising. The red channel is generally the most troublesome in this respect, and for this reason, the red-beam splitter is commonly placed first in the light path. Fig. 1 shows it to occupy the second position in the present scanner. It would appear that the combined effect of the light loss brought about by the arrangement adopted for the raster-brightness servo mechanism (Fig. 9) and the placing of the red-beam splitter is to lower the signal/noise ratio in the red channel by about 3 dB. Alternative techniques might substantially reduce this figure.

THE AUTHORS' REPLY TO THE ABOVE DISCUSSION

Messrs. H. E. Holman, G. C. Newton, and **S. F. Quinn** (*in reply*): We agree with Mr. Gouriet's suggestion that colour analysis is another example of the principle by which information can be exchanged for signal/noise ratio, and there is no doubt that the transmission of redundant information reduces the signal/noise ratio. XYZ co-ordinates are naturally inefficient, since the area of natural colours in the chromaticity diagram occupies only a small part of the XYZ triangle. The redundant information is therefore considerable, and this leads to a worse signal/noise ratio than with the RGB system. We confirm Mr. Jesty's view that other possible analysing primaries give correct colour analysis and better signal/noise ratios than the XYZ primaries, but we consider it doubtful whether the colour rendition of commercial film justifies such perfection.

In reply to Dr. Rendall, the geometrical and shift waveform accuracies are maintained to within a quarter of a line, which increases towards the extreme corners where an error of one line is found. This may be in both the x and y directions.

The accuracy of patch registration is such that, when using a black-and-white test chart and observing on a three-tube monitor, 500 lines per picture width can be resolved within a circle whose diameter is 0·8 of the height of the image. The pre-focus accuracy of location depends upon precise tube manufacture, and it is possible, within five minutes, for a tube to be replaced by a semi-skilled operator and to be providing a picture with the same standard of resolution.

The problem of stability of screen brightness with feedback depends mainly upon the uniformity of the semi-reflecting mirror surface. In practice, variations of less than 1% have been achieved over the effective area.

Tube life has been limited by screen ageing due to electron burning, but glasses are available which offer a life of many hundred hours.

With regard to Mr. Vinten's observations, it should be mentioned that there are significant differences in the behaviour of good- and bad-quality 16mm film. Until the position of a taking frame relative to its locating perforation is standardized by film-camera manufacturers, and this relationship is upheld by other process- and projection-equipment manufacturers, random errors are bound to occur.

In reply to Mr. Jacobs, this beam-splitter arrangement was chosen to give the simplest spectral characteristics to the dichroic mirrors. The red portion of the spectrum is attenuated by 6% in the first dichroic mirror, decreasing the red signal/noise ratio by 0·3 dB, which is negligible. The screen-brightness feedback increases noise by reducing the light available (approximately 2 dB), and also by adding another noise source (approximately 3 dB), but it does not affect the relative signal/noise ratios of the three colour components. This increase of 5 dB, however, represents the difference between a very satisfactory noise level and one just satisfactory on 16mm film.

THE APPLICATION OF DIGITAL COMPUTERS TO NUCLEAR-REACTOR DESIGN

By J. HOWLETT, B.Sc., Ph.D., Associate Member.

SUMMARY

The paper reviews the main computational problems arising in the design of a nuclear power reactor. The numerical-mathematical methods available are described briefly in two broad classes, namely the discrete-particle (Monte Carlo) treatment of the neutron-transport problems and the analytical methods based on the transport equation. The use of digital computers in the different methods is discussed in general terms and three examples are considered in more detail: the calculation of the thermal utilization factor for a lattice by the spherical harmonics method, the direct numerical solution of the 2-group diffusion equations for a cylindrical reactor and the calculation by Monte Carlo methods of the resonance escape probability for a graphite-uranium lattice. Finally there is a discussion of the standard and performance of computers likely to be required by nuclear engineers.

LIST OF PRINCIPAL SYMBOLS

ϕ = Neutron flux (suffixes f and s for fast and thermal fluxes, respectively), neutrons/cm^2/sec.

η = Number of secondary neutrons produced per capture.

ε = Fast fission factor.

p = Resonance escape probability.

f = Thermal utilization factor.

α = Inverse of mean free path for a neutron, cm^{-1}.

β = Number of secondaries per unit path, cm^{-1}.

Σ_{af}, Σ_{am} = Macroscopic absorption cross-sections of fuel and moderator, respectively, cm^2.

L_s = Slowing-down length for fast neutrons, cm.

L = Diffusion length for thermal neutrons, cm.

D_f, D_s = Flux diffusion coefficients for fast and thermal neutrons respectively, cm.

r, z, θ = Radial, axial and angular co-ordinates; $\mu = \cos\theta$.

n, j, k (used as suffixes) = Integers.

(1) THE SCOPE FOR COMPUTER APPLICATIONS

The computational effort involved in the design of a nuclear power reactor is much greater than in any other engineering enterprise, principally for the following reasons:

(a) The capital costs of a reactor and its fuel charge are so great that the design must be most carefully optimized and therefore examined in great detail.

(b) The possible consequences of a design error are so serious that the performance must be predicted as accurately as possible.

(c) The nuclear events occurring in the reactor are complex and the equations representing them cannot be handled in any simple way; also the basic data (mostly nuclear cross-sections) have complicated dependences on energy, not representable by any simple mathematical forms.

(d) Much of the information needed in a design study is obtained by analysis of 'integral' (e.g. exponential) experiments; this analysis is a considerable computational problem and needs to be done quickly to be of value.

(e) Large-scale experiments, such as zero-energy assemblies, are serious undertakings which can be made only after most of the design questions have been settled by calculation.

To see how and where in a design study the computing problems arise, let us consider the main steps in the process, assuming that the reactor is intended to drive an electrical power station.

First, a decision must be taken, based on very broad economic and scientific arguments, on the electrical output required and the type of reactor to be built. Thus, if 100 MW of electrical energy is the aim, the reactor will have to deliver about 400 MW of heat; if it is to be a natural or slightly enriched uranium thermal reactor (as is most likely for several years to come) the fuel rating will be about 4 MW per ton and the charge will therefore be about 100 tons. The moderating ratio will have to lie in the range from about 1 : 1 to 5 : 1, so the total weight of moderator will lie between about 100 and 500 tons; the shape is not likely to differ greatly from a right cylinder, so the dimensions can be quickly settled within reasonable limits.

General engineering and metallurgical considerations will determine a range for the working temperature, and the known problems of pumping—the power required and possible corrosion effects—will settle general questions on the circulation of the coolant, which may be the moderator also, and will thus provide some criteria for the design of the heat exchanger.

The main programme of nuclear calculations can now start. The aim will be to find the dependence of the critical mass of the core on the moderating ratio, enrichment of fuel, temperature and dimensions, all these quantities being allowed to vary over the ranges suggested by the preceding general arguments. The calculations will give in each case the flux distribution and hence the fission rate (which determines the power level) and the conversion factor for the build-up of fuels such as plutonium 239 or uranium 233 if this is a relevant consideration. These criticality calculations will use such integral nuclear properties of the core design as thermal utilization factor, resonance escape probability and slowing-down and diffusion lengths, which must themselves be calculated from design details and basic nuclear data.

Heat-transfer calculations will show how the supply of energy to the electrical station varies with the reactor coolant circulation rate and outlet temperature (and pressure if it is a gas); these will probably have to be done for several sets of design details of the heat exchanger. Much of this depends upon empirical laws.

If these calculations have been made in sufficient detail, there will be enough information to settle within closer limits the essential details of the design on economic grounds, balancing the value of the electrical output and any fuel breeding against the cost of the reactor, fuel charge, heat exchanger and generating equipment.

With the narrowed ranges for the design parameters, further calculations will be needed, to give

(a) The kinetics of the reactor, i.e. the response to movements of the control elements, particularly when starting up or shutting down; the possibility of dangerous surges in power or temperature must be investigated.

(b) Long-term changes in reactivity due to the accumulation of fission-product poisons, to fuel breeding and to burn-up of the fuel charge; the results here will enable a fuel-cycling schedule to be worked out.

Dr. Howlett is at the U.K.A.E.A. Atomic Energy Research Establishment.

(c) The shielding requirements; there are worth-while economies to be achieved by avoiding excessive amounts of shielding beyond what is needed for safety.

The results of this last group of calculations will probably lead to a re-examination of the earlier optimization, and possibly to a need for more criticality calculations; after a few repetitions of the whole process it should be possible to settle definite values for the design parameters.

Almost all the separate calculations here are complicated and of large scale; it is not possible to work through such a programme in a reasonable time, and in the detail required, without powerful computing machinery. It would be possible to build a special machine to perform any one of the calculations, as has, in fact, been done,[1-3] but because the various types are so different mathematically, any organization which has to carry out complete studies will need to have—or at least to have access to—a fast general-purpose digital computer. The field is an ideal one for the application of these machines: the amount of computation to be done is very large, so the benefits of high speed are great; there is a wide variety of problems, so that the machine's flexibility is exploited; much of the work has to be done to high numerical precision, so good use is made of the ability to carry many figures in the calculation; and all the calculations have to be repeated many times with different values for the parameters, so a good return is obtained for the effort put into the development of the machine programmes. In this connection there is one very important advantage of machine computation: if a large survey has to be made with real urgency, and if it is important enough to command exclusive use of the machine, then it can be carried through simply by continuous running, all the staff available taking turns to operate the machine.

Without a machine a design team must prune its demands very drastically, both by making do with a much less detailed survey than it would like to have and by using methods known to be of lower accuracy than others which, though preferable, would use more computing time, such as using diffusion-theory approximations when these are not justified or treating as spherically symmetrical a system for which a finite cylinder is a much better representation. This question of approximations certainly arises in computer applications, but less sharply and at a less elementary level. Of course, reactors have been designed successfully without the aid of fast computers, but these were essentially non-commercial plants where economic questions hardly arose; for full-scale industrial use reactors must be designed with close regard to these questions, and they cannot be settled without extensive and detailed calculations.

(2) MATHEMATICAL AND COMPUTATIONAL METHODS AVAILABLE

Of the essential computational problems of reactor design just listed the largest and most characteristic are those concerning the nuclear properties, which involve a study of the neutron population. This is necessarily a difficult mathematical problem; a neutron is described by seven co-ordinates: three for space, two for angle (for the direction of motion), speed (or energy) and time; the important events in its life are collisions with nuclei, the results of which are described by complicated laws deduced from fundamental nuclear-physics experiments. The methods available for attacking these problems of neutron transport can be grouped as follows:

(a) Exact formulation in terms of the laws which describe the movements of individual neutrons; these state the probabilities of various events, such as collision, absorption, fission. The method is statistical and, because of its use of random numbers in the calculation, is usually known as the Monte Carlo method.

(b) Idealized formulation, regarding the neutron population

as continuous and setting up a conservation equation as in hydrodynamics; this is an integro-differential equation, known as the Boltzmann transport equation. In certain circumstances, e.g. in systems large compared with the neutron mean free path (and therefore in much of the work on thermal reactors), it can be reduced to the partial differential equations of the diffusion-theory approximation. From either of these equations one can proceed in either of two ways:

(i) Attempts at formal analytical solution, e.g. by expansion in series or by integral (e.g. Laplace) transforms. An example is the spherical harmonics method[4] for solving the transport equation; a standard process for solving the diffusion equations in a cylindrical system is to expand in series of Bessel functions.

(ii) Direct numerical solution, without recourse to analysis; all differential coefficients are replaced by differences and all integrals by sums, so that the problem is reduced to that of solving a system of algebraic equations. Section 3.2 shows how this is applied to the diffusion equations; Carlson[4,5] has developed a corresponding method for solution of the transport equation.

Examples of uses of digital machines in these contexts are given later; it seems suitable here to give some general consideration to the different methods.

The application of a fast automatic computer to a mathematical problem can be approached in two ways. One can either attempt a solution by formal mathematical analysis, carrying this as far as possible and bringing the machine to evaluate the expressions obtained, much as one would use hand-computing methods; or one can develop a method which is numerical from the start and which is designed with the machine's properties in mind. This is paralleled above in the division into methods (i) on the one hand and (a) and (ii) on the other. The analytical method has the advantage that the work is kept general until a late stage and numbers are inserted only when analysis can go no further; the final computations usually go through quickly (see Section 3.1) even though they may seem very heavy for hand computation; thus a large parameter survey can be made in a reasonable time. The disadvantages in reactor work are all consequences of the complications of the mathematics. First, only simple geometries can be dealt with; 1-dimensional systems are reasonably tractable (e.g. spheres or infinite cylinders), 2-dimensional systems are much more difficult (e.g. finite cylinders with symmetry about the axis) and 3-dimensional systems are quite unmanageable. Secondly, only simple energy-dependence can be allowed; most of the analytical treatments have been made with at most two energy groups, and some work has been done with simple functional energy dependence, leading to heavy algebra,[6] but no treatment of detailed energy dependence seems possible. Finally, almost all such methods are approximate in the sense that they depend upon expansions into series with a limited, usually very small, number of terms retained; any attempt to improve on this is likely to be defeated by the weight of the algebra. Thus, in the spherical harmonics method the standard practice is to go as far as the P_3 term; the next approximation, taking P_5, requires much more labour and it is doubtful whether anyone has ever gone further. All this is not to be taken to imply that these analytical methods are of no value; on the contrary, they provide an exceedingly valuable way of surveying a wide field quite quickly as a first approximation, and so narrowing the ranges over which more accurate methods have to be used. But the advent of fast automatic machines has reduced the incentive to press this treatment further, and it is unlikely that there will be any important new developments; thus Davison's book[7] can probably be taken as the final summing-up of the analytical theory of neutron transport.

The purely numerical methods [(a) and (ii) above] are much more in the spirit of automatic computers and could not be used effectively until such machines had been developed; both are likely to involve large—probably very large—numbers of

rithmetical operations, but most of these will be arranged in airly simple patterns repeated many times, which is the ideal situation for computer programming.

With method (ii), since analytical difficulties are avoided, more complicated situations can be considered, such as genuinely 2-dimensional systems and detailed variations of nuclear properties with energy; but the time required to achieve a solution will increase with the complication of the problem, so there are practical limits to the applications. For example, 3-dimensional problems can, in principle, be solved by these methods, but are, in fact, still unmanageable, simply because even the fastest of present-day machines is too slow. In place of the analytical difficulties, one is now faced with questions of convergence and stability of iterative processes and of the effects of truncation errors (i.e. of the approximations introduced when differential coefficients or integrals are replaced by finite differences or sums); these are important and difficult questions—more exacting, in fact, than the technical complications of heavy algebra—and give plenty of scope for mathematical study.[8-11] The effects of truncation errors can usually be investigated empirically by making exploratory calculations (on the machine, of course) with different intervals in the differencing process, or different forms of approximation; slowness of convergence and instability (i.e. the computed solution running away from the true solution, owing to build-up of rounding errors) are usually made obvious by the progress of the calculation but may be very difficult to overcome. An unstable computing process is, of course, useless; an improvement in the rate of convergence may, by greatly reducing the time taken to find an acceptable solution, make all the difference between an academic mathematical process and a practical computational method; Section 3.2 illustrates this. The effect of truncation errors is seen very clearly in one important reactor calculation; an attempt to integrate the equations for the reactor kinetics[12] by a step-by-step method using the standard Runge–Kutta process[13] may prove impractical because of the need to use extremely small steps, and therefore very long machine runs, even though the neutron densities are changing very slowly with time; this is due to the fact that the method is equivalent to taking at each step a fixed number (4 is usual) of terms of an expansion for the solution in powers of the interval, and the numerical values may be such that this is a poor approximation unless the interval is very small. The only remedy is to use a method having a different mathematical basis, such as one of the Fox–Goodwin processes (see Reference 14, which also refers to this particular difficulty).

The Monte Carlo method is the most general possible and can, in principle, be used to solve any reactor problem, however complex. The essential features are the tracing, by calculations using random numbers, of the life histories of samples from the neutron population and the inferring from these of the properties of the population as a whole. Because it abandons analysis completely, even in the formulation of the problem, this method provides the best hope for dealing with complicated geometry or energy dependence. It also provides the only means of accurate formulation of some problems; for example, the conditions at the surface of a hole inside a reactor can be expressed only approximately in diffusion theory, and in terms of transport theory are complicated and very difficult to use, but the Monte Carlo method uses directly the simple physical fact that a neutron emerging into a hole will travel straight across at constant speed. The drawback to the method is that, being statistical, it gives results in the form, 'the probability is P that the parameter k lies between a and b', and for good accuracy, i.e. the location of the number required between narrow limits with high probability, a very large amount of computation may be required. However, considerable gains can be made by proper design of the calcula-tion (which can be regarded as a theoretical experiment), and recent developments in statistical techniques can give large improvements in the efficiency.[15-17]

(3) EXAMPLES OF MACHINE APPLICATIONS

During the last few years much effort has been put into programming reactor calculations, particularly in the United States, where large and fast machines have been much more readily available than in Britain. The mere cataloguing of the published information on these programmes is already a formidable undertaking; summaries have been issued by Radkowsky and Brodsky[18] and by the U.S.A.E.C. Computing Facility at New York University,[19] and several comments and references are given by Sangren.[20]

It is, of course, possible to write a machine programme to perform any calculation, and with modern machines one can usually expect to get an increase in speed over hand computation of at least 1 000; whether in any particular case this is a reasonable or economic thing to do is an entirely different question. All the standard reactor calculations have been programmed, most of them for all makes of commercially produced machine both here and in America. Many of the American computing laboratories have produced very general large-scale programmes, e.g. the 'Eyewash' programme[21] for solving the multi-group, multi-region 1-dimensional diffusion equations with Univac, and the 'Cure' programme for the multi-group multi-region 2-dimensional diffusion equations with the IBM 704. In England much work has been done with the Deuce at the National Physical Laboratory on the integration of the reactor-kinetic equations. The examples which follow are for important general calculations of which the author has had direct experience; they are given in rather general terms, for it seems inappropriate either to quote large amounts of algebra or to reproduce programme details. It is hoped that, while indicating the powers of the machine, they also show the importance of a sound knowledge of numerical mathematical technique to anyone who must do a significant amount of machine programming.

(3.1) The Spherical Harmonics Method

This is an analytical process for solving the transport equation whose practical application is limited to 1-dimensional systems with one energy group. One of its most frequent applications is to the calculation of the thermal utilization factor of the lattice of a heterogeneous thermal reactor; diffusion theory is not a good approximation here, because, although the reactor as a whole is large compared with the neutron mean free path, the dimensions of interest are the diameters of the rods and cooling channels, which are comparable with the mean free path.

In a spherically-symmetrical system with one energy group the transport equation can be written

$$\mu \frac{\partial \Phi}{\partial r} + \frac{1 - \mu^2}{r} \frac{\partial \Phi}{\partial \mu} + \alpha \Phi = \tfrac{1}{2}\beta \int_{-1}^{1} \Phi(r, \mu') d\mu' \quad . \quad (1)$$

where $\Phi(r, \mu)$ is the flux at radius r for neutrons whose direction of motion is specified by $\theta = \arccos \mu$. The essence of the method is to assume a solution

$$\Phi(r, \mu) = \tfrac{1}{2} \sum_{n=0}^{\infty} (2n + 1) P_n(\mu) \phi_n(r) \quad . \quad . \quad . \quad (2)$$

where $P_n(\mu)$ is the nth Legendre polynomial (spherical harmonic) and ϕ_n are functions of r to be determined. Substitution in the equation and elimination of μ leads to a set of ordinary differential equations for ϕ_n, which can be solved formally in terms of Bessel

functions. The constants entering into these solutions are found by fitting the general solution to the given boundary conditions, which express, for example, the fact that at the outer boundary no neutrons enter the system; this part of the calculation requires the setting-up and solving of a set of linear equations, formally of infinite order but made finite in practice by taking only a finite number of terms in the series (2). Thus, an approximation to the solution is found in terms of known functions; integration over all values of μ gives the scalar flux from which the thermal utilization factor, f, can be found by integration, i.e.

$$f = \int \Sigma_{af}\phi dV \Big/ \int (\Sigma_{af} + \Sigma_{am})\phi dV \quad . \quad . \quad . \quad (3)$$

where Σ_{af} and Σ_{am} are the absorption cross-sections for the fuel and moderator respectively and the integrals are taken over the volume of a lattice cell; these integrals can be evaluated analytically and an explicit expression for f can be found.

The analysis has been generalized by Tait[22] for cylindrical geometry, the algebra being a good deal more complicated than that for spherical symmetry. This paper gives the complete algebraic treatment for the P_3 approximation (i.e. taking only the terms P_1 and P_3 in the expansion), which uses Bessel functions up to I_3 and K_3, and requires the solution of a set of nine linear simultaneous equations. This has been programmed for a commercial machine by Preston,[23] the calculation taking 4 min for the f-value and—if these are required—a further 4 min for the fluxes at 10 radial values; hand calculation takes a full week. The programme is elaborate, with about 3 000 machine instructions; it is fully automatic, computing all the Bessel and other functions required and setting up and solving the linear equations, and is written in floating-point arithmetic to give flexibility in the sizes of the physical data which may be inserted. The programme shows very clearly the power of the machine (not even a particularly fast one in this case) in dealing with a complicated algebraic process.

(3.2) Critical Size of a Cylindrical Reactor: Direct Numerical Solution of the 2-Group Diffusion Equations

The problem here is to find the criticality conditions for a reacting system in the form of a finite cylinder symmetrical about its axis; the simplest case is that of a cylindrical homogeneous core surrounded by a homogeneous cylindrical reflector, all dimensions being given and the critical value of the reactivity being required. It is assumed that system is large and has no internal cavities, so that diffusion theory can be used.

If two energy groups of neutrons are considered (fast and thermal) with fluxes ϕ_f, ϕ_s respectively, the equations are

$$\left.\begin{array}{l} D_f(\nabla^2 - 1/L_s^2)\phi_f + \eta\varepsilon f D_s\phi_s/L^2 = 0 \\ D_s(\nabla^2 - 1/L^2)\phi_s + p D_f\phi_f/L_s^2 = 0 \end{array}\right\} \quad . \quad . \quad (4)$$

where D_f and D are the two diffusion coefficients, and L_s and L are the slowing-down and diffusion lengths. The operator ∇^2 is here

$$\partial^2/\partial r^2 + 1/r\partial/\partial r + \partial^2/\partial z^2$$

The usual boundary conditions are that the fluxes vanish over the outer boundary and the fluxes and neutron currents are continuous at any internal boundary between two different media. The mathematical problem is an eigenvalue one, i.e. to find the value of the product $\eta\varepsilon f$ in the core for which the equations have a solution which is not everywhere zero.

The basis of the direct method of solution is the replacement of the differential coefficients by finite differences. If a flux ϕ depends on a single variable r, and if we take a set of equally-

spaced points $r_0, r_1 \ldots r_n$ with distance apart h, then approximations to the derivatives are

$$\left.\begin{array}{l} (d^2\phi/dr^2)_{r=r_n} = \dfrac{1}{h^2}(\phi_{n+1} - 2\phi_n + \phi_{n-1}) \\[2mm] (d\phi/dr)_{r=r_n} \doteq \dfrac{1}{2h}(\phi_{n+1} - \phi_{n-1}) \end{array}\right\} \quad . \quad . \quad (5)$$

both correct to order h^2. Thus the differential equation

$$d^2\phi/dr^2 + a^2\phi = 0 \quad . \quad . \quad . \quad . \quad (6)$$

can be replaced approximately by the set of algebraic equations

$$\phi_{n+1} - (2 - a^2h^2)\phi_n + \phi_{n-1} = 0 \quad . \quad . \quad (7)$$

If the equation is to be solved with the boundary conditions $\phi = 0$ at $r = 0$, and $\phi = 1$ at $r = L$, so that $r_n = L$ and $h = L/n$, there is one equation of the type (7) corresponding to each of the $n - 1$ internal points; together with the known values of ϕ_0 and ϕ_n, these suffice to determine the $n - 1$ unknown values of ϕ and the solution gives an approximation to the solution of the original differential equation. Clearly the accuracy of the approximation will increase as the number of points is increased, but so will the amount of work involved in finding the solution; also there are mathematical objections to taking excessively small intervals.

In the 2-dimensional problem we divide the region by mesh-lines parallel to the axes and consider the fluxes at intersections of these lines. At the mesh-point $r_{jk} = (j\delta r, k\delta z)$ the approximation to the term $\nabla^2\phi$ is now

$$\frac{1}{(\delta r)^2}(\phi_{j+1,k} - 2\phi_{j,k} + \phi_{j-1,k}) + \frac{1}{2r_j(\delta r)}(\phi_{j+1,k} - \phi_{j-1,k})$$
$$+ \frac{1}{(\delta z)^2}(\phi_{j,k+1} - 2\phi_{j,k} + \phi_{j,k-1}) \quad (8)$$

Using this we can replace each partial differential equation by an approximation consisting of a set of linear algebraic equations, one for each mesh-point, each involving the fluxes at that point and a few of those surrounding it. The mathematical problem of solving these is simple in principle but is difficult in practice, because of the size of the task; one must take about 1 000 points to give a reasonable representation of an actual reactor, and the problem of solving a set of 1 000 linear equations, even of the special form of those in question, is a serious one.

Hassitt has written a programme[24] for a commercial machine which will solve this problem for a symmetrical finite cylindrical reactor, allowing almost any variation of properties from point to point; in particular, any system of cylindrical shells, with top and bottom reflectors, can be dealt with. The programme solves the criticality problem; thus, for the simple case of a uniform core surrounded by a uniform reflector, all of given dimensions, it will produce the value of $\eta\varepsilon f$ giving a critical system, and the corresponding flux distributions. The process of solution is iterative, the programme calculating a sequence of approximate fluxes which converges to the solution of the algebraic equations (which, of course, is itself an approximation to the actual problem) and printing out after each iteration the value of the current approximation to the criticality parameter. The user of the programme can see how these values are converging and must decide when convergence has gone far enough, when he can stop the iteration and print out the tables of fluxes. Apart from this the programme is completely automatic.

This is a large-scale calculation, because with the large number of mesh points usually required the iteration converges slowly and eight or more complete cycles may be needed to give the

value of the parameter correct to four significant figures; as with all eigenvalue problems the convergence of the fluxes is slower. It is not at all unusual for a calculation to take two hours on the commercial machine, which means that a parameter survey is likely to be a serious undertaking, really too large a task for a machine of this speed. Since it is also a quite fundamental calculation in thermal reactor design and one which must be carried out many times in any design study, it shows very well the need for large and fast machines. High arithmetical speeds are needed because, even in a simple problem of this type, several million operations must be performed; but because the calculation uses a large field of numbers, a machine with a small fast store will use much time transferring numbers between this and the low-speed backing store, and therefore a large fast store is essential if the complete calculation is to be done quickly. The commercial machine referred to above has a fast store for 384 numbers, which is small on this scale; a development of it, with arithmetical speeds about 10 times those of the first and a fast store of 1 024 numbers, would probably give an increase in overall speed of about 5–10, while a third machine, which has arithmetical speeds similar to those of the second but can be had with very large fast stores (up to 32 000 numbers) probably gives an improvement of 10–20.

This example illustrates very well the simplicity of the direct numerical approach and the scale of the computing problem to which it leads; hand calculation is here out of the question. Any attempt at analytical treatment of this problem, unless it makes drastic simplifications, leads to excessively heavy algebra; furthermore, the direct method is easily extended to deal with more than two energy groups—at a cost of longer calculations, of course—while the labour involved in doing this analytically would be prohibitive. One important point of detail here is that the iterative process used in solving the finite-difference equations must be designed with great care; the general idea of solving partial differential equations in this way has been known for a long time, but it was also known that a straightforward iterative method converged too slowly to be of use, even with a fast machine. It was not until the essential mathematics of the process had been studied in detail, notably by Frankel[25] and Young[26], that a means of accelerating the convergence was found which led to the possibility of practical application on a large scale.

5.3) Monte Carlo Calculation of the Resonance Escape Probability for a Lattice

The resonance escape probability, p, for a reacting system is one of its most important properties from the aspect of the neutron economy, for it enters into the '4-factor' formula for the (infinite) multiplication factor

$$k = \eta \varepsilon p,$$

η is a fundamental nuclear property, found by experiment, whilst ε, p and f are derived quantities depending on the geometry of the system as well as on the nuclear properties of the materials. A method for calculating f has been described in Section 3.1; calculation of p is much more difficult, because its value will depend upon the whole energy spectrum of the neutron population from fission to thermal and upon the variation of the absorption and scattering cross-sections over this range. The value will vary with the temperature of the reactor, because of the broadening of the resonances in the absorption cross-section with increasing temperature; this effect must be taken into account in a design study, and is particularly difficult to measure experimentally. Calculation by an analytical method seems to be out of the question, since this would require the solution of the full transport equation in three dimensions and with the

actual variation of cross-sections with energy; thus the only possibility is a Monte Carlo approach. Much work on these lines has been done by Richtmyer;[27] the only similar English work known to the author is that done by K. W. Morton of Harwell.

Morton studied the slowing-down of neutrons in a lattice of cylindrical uranium rods, each surrounded by an empty channel, in a graphite moderator; as a first step he took the range from 50 eV to thermal, which includes three of the strongest resonances. The essence of the calculation is that a neutron is started from some point in the system with energy near the top of the range and is allowed to migrate, undergoing collisions with fuel and moderator nuclei until it is either absorbed or reduced to thermal energy, the starting conditions, the positions of the successive collisions and the results of these being decided by choices of random numbers. The tracking is repeated with more neutrons until enough have been taken to give an estimate of the escape probability of the required accuracy. Thus if N neutrons are started and n survive to thermal energy, an estimate of p is n/N; but this has a standard deviation, which measures the uncertainty of the estimate, of $\sqrt{[p(1-p)/N]}$. What is wanted is an estimate reliable to a few parts in 1 000, so the number, N, of neutrons to be taken is found from

$$\frac{1}{p}\sqrt{\left[\frac{p(1-p)}{N}\right]} = \frac{m}{1\,000}$$

where m is a small integer. In fact, p is usually about 0.9, so N must be about 10 000.

In his preliminary calculations Morton took samples of 4 000 neutrons for each of four temperatures, and by use of various refinements in the statistical analysis obtained accuracy of 2% in the values of $1 - p$; the running time on the computer was 25 min for each case. Each calculation used about 25 000 random numbers, which were computed by the programme itself; the effect on the cross-sections of the thermal motion of the scattering nuclei was also computed by the programme, using the 'ψ-function',

$$\psi(x, t) = \frac{1}{2\sqrt{(\pi t)}} \int_{-\infty}^{\infty} \varepsilon^{-(x-y)^2/4t} \frac{dy}{1 + y^2}$$

where x is related to energy and t to temperature; it is an interesting comment on the speed of the computer used that, although extensive tables of this function have been published and could have been supplied as input, it was found preferable to programme an independent calculation and compute the function afresh at the beginning of each case.

(4) MACHINE DEVELOPMENTS IN THE FUTURE

The examples just given, which are typical of the analytical, direct numerical and Monte Carlo approaches to reactor calculations, indicate the extent to which machines of various standards of performance meet the demands made on them. The Mark I* computer, with a speed of about 500 arithmetical operations per second, is quite adequate for even a complicated algebraic process but is about at its limit in the solution of the 2-dimensional criticality problem: machine runs of two hours and more per solution are uncomfortably long, especially when a failure near the end may vitiate all the preceding work; also such long times make it difficult to run through a parameter survey (say of 10–20 solutions) in any reasonable time unless the work is given high priority. The Mercury and the 704 computers, with speeds about 10–20 times that of the Mark I*, are adequate for this problem, the latter gaining an extra advantage from its larger fast store. Similar remarks apply to the Monte Carlo calculation.

However, it is certain that as reactor design standards advance, engineers will want more elaborate and extensive calculations which will be beyond the scope of even the best of present-day machines. The criticality calculation for a genuinely 3-dimensional system is an example. In the 2-dimensional programme described the number of flux values stored is between 10^3 and 2×10^3 and the number of arithmetical operations performed is about 10^6. The 3-dimensional calculation would require more flux values, more operations at each mesh point in the course of each iteration and more iterations, because the convergence would be slower; making reasonable estimates for these effects, one is led to the conclusions that the number of fluxes to be stored is of the order of 10^5 and the number of operations at least 10^8, even if one uses only a 2-group approximation. Since such calculations would be wanted in batches, either for parameter surveys or in theoretical studies of the life history of a proposed reactor, a running time of about 15 min (say 10^3 sec) would be desirable; thus the machine speed would have to be at least 10^5 operations per second, i.e. at most 10 microsec per operation. Furthermore, to avoid serious loss of overall computing speed in transfers between the fast and backing stores, it would be necessary to have a fast store able to hold most of the number field, i.e. for about 10^5 numbers. It is open to question whether one would, in fact, use simple diffusion theory for a calculation on this scale; it is not unlikely that a Monte Carlo method would be used, to permit a more realistic representation of the reactor structure. But what is certain is that the standards set by the needs of the diffusion-theory calculation will not be higher than those for more accurate methods, so we may take these as the minimum to be aimed at.

It is known that there are several projects under way in America for building machines of at least this standard. The two about which most information is available are the Larc and the Stretch, both sponsored by the United States Atomic Energy Commission. The first will run at about 10^5 operations per second with fast storage for some multiple of 10^4 numbers—figures of 50–60 000 have been quoted—and is due for completion in 1958. The corresponding figures for the other are 10^6 operations per second and 10^5 numbers at least, this machine being due for completion in 1960. The name Stretch was given to bring out the fact that this speed is approaching the limit which can be reached with known or foreseeable components; for instance, a time of 1 microsec for a complete multiplication requires that the circuit elements (transistors) shall operate at 100–200 Mc/s, and at this frequency the time taken for the pulses to travel round the computer is beginning to limit its physical size. It is thus possible that this order of performance may represent the end-point of computer development, apart from such arrangements as parallel operation of arithmetic units or even complete machines to give increased effective speed; if this proves to be so, the development will have been taken from start to finish in the short time of about 25 years, and undoubtedly the needs of the nuclear-energy world will have provided an important part of the stimulus.

(5) ACKNOWLEDGMENT

(6) REFERENCES

(1) WILSON, I., and BURNUP, T.: 'An Analogue Computer for Reactor Kinetic Equations', Atomic Energy Research Establishment, Harwell, Report AERE RE/R 1485, September, 1954.
(2) LIEBMANN, G.: 'Solution of Some Nuclear Reactor Problems by the Resistance Network Method', Journal of Nuclear Energy, 1956, 2, p. 213.
(3) BAYLY, G., and PEARCE, R. M.: 'Method of Studying Multi-Region Reactors with an Analogue Computer', Atomic Energy of Canada, Ltd., Chalk River Report CRRP–606, April, 1956.
(4) CODD, J., SHEPHERD, L. R., and TAIT, J. H.: 'The Physics of Fast Reactors', Progress in Nuclear Energy, 1956, 1, p. 251.
(5) CARLSON, B. G.: 'Solution of the Transport Equation by the S_n Method', Los Alamos Scientific Laboratory of the University of California, Report LA–1891, April, 1955.
(6) ELLIOTT, J. P.: 'A Method of Solving the Velocity-Dependent Boltzmann Equation', Atomic Energy Research Establishment, Harwell, Report AERE T/R 1429.
(7) DAVISON, B.: 'Theory of Neutron Transport' (University Press, Oxford, 1957).
(8) HOUSEHOLDER, A. S.: 'Principles of Numerical Analysis' (McGraw-Hill, 1953).
(9) HOUSEHOLDER, A. S.: 'Generation of Error in Digital Computation', Oak Ridge National Laboratory, Report ORNL–1938, October, 1955.
(10) CURTISS, J. H. (Ed.): 'Numerical Analysis—Proceedings of the Sixth Symposium in Applied Mathematics of the American Mathematical Society', Santa Monica, August, 1953 (McGraw-Hill, 1956).
(11) O'BRIEN, G. G., HYMAN, M. A., and KAPLAN, S.: 'Numerical Solution of Partial Differential Equations', Journal of Mathematics and Physics, 1951, 29, p. 223.
(12) GOERTZAL, G.: 'Reactor Dynamics', Chapter 1.6 of 'Reactor Handbook, Vol. 1', United States Atomic Energy Commission, AECD–3645, 1955.
(13) WILLERS, F. A.: 'Practical Analysis' (Dover, 1948), Chapter 6.
(14) FOX, L., and GOODWIN, E. T.: 'New Methods for Numerical Integration of Ordinary Differential Equations', Proceedings of the Cambridge Philosophical Society, 1949, 45, p. 373.
(15) KAHN, H.: 'Applications of Monte Carlo', United States Atomic Energy Commission, AECU–3259, April, 1954.
(16) MORTON, K. W., and HAMMERSLEY, J. M.: 'Antithetic Variates, a new Monte Carlo Technique', Proceedings of the Cambridge Philosophical Society, 1956, 52, p. 449.
(17) MORTON, K. W.: 'Criticality Calculations by Monte Carlo', Atomic Energy Research Establishment, Harwell, Report AERE T/R 1903.
(18) RADKOWSKY, A., and BRODSKY, R.: 'A Bibliography of Available Digital Computer Codes for Nuclear Reactor Problems', United States Atomic Energy Commission, AECU–3078, 1955.
(19) NUCLEAR CODES GROUP of the Atomic Energy Commission Computing Facility, Institute of Mathematical Sciences, New York University.
(20) SANGREN, W. C.: 'The Role of Digital Computers in Nuclear Design', Nucleonics, May, 1957, 15, p. 56.
(21) ALEXANDER, J. H., and GIVEN, N. D.: 'A Machine Multi-Group Calculation', Oak Ridge National Laboratory, Report ORNL–1925, 1955.
(22) TAIT, J. H.: 'Calculation of the Thermal Fine Structure in a Pile by the Spherical Harmonics Method', Atomic Energy Research Establishment, Harwell, Report AERE T/R 371, 1949.
(23) PRESTON, P. D.: 'The P3 Program', Atomic Energy Research Establishment, Harwell, Memorandum T/M 147, 1957.
(24) HASSITT, A.: 'A Program for Solving the Two-Group Diffusion Equations with a Digital Computer', Atomic Energy Research Establishment, Harwell, Memorandum T/M 148, 1957.
(25) FRANKEL, S.: 'Convergence Rates of Iterative Treatment of Partial Differential Equations', Mathematical Tables and other Aids to Computation, 1950, 4, p. 65.
(26) YOUNG, D. M.: 'On the Solution of Linear Systems by Iteration', p. 283 of Reference 10.
(27) RICHTMEYER, R. D.: 'Resonance Capture Calculations in Lattices by Monte Carlo Methods', Report on Conference on Resonance Absorption, Brookhaven National Laboratory, September, 1956, BNL 433.

TEMPERATURE TRANSIENTS IN GAS-COOLED THERMAL NUCLEAR REACTORS

By J. H. BOWEN, B.Sc., and E. F. O. MASTERS, B.Sc., Associate Members.

SUMMARY

The paper examines the transient behaviour of the Calder Hall type of reactor in terms of the design and operating parameters, the main transients examined being those arising from

(a) Step changes in reactivity.

(b) Steady changes in reactivity including the start-up condition.

(c) Step changes in cooling.

(d) Steady and other changes in cooling due to both circulator-speed and gas-pressure changes.

The examination has been carried out by means of an analogue computer and by a graphical method which employs an approximate solution of the neutron-kinetic equations. The results obtained by the two methods are compared and agreement is shown to be good.

The graphical method has been extended to take account of temperature gradients within the fuel elements and fission-product heating. Account is also taken in the computations of the heat generated in the graphite.

The effects of non-uniformity in the axial and radial temperature distribution on the temperature transient throughout the reactor are examined and assessed.

It is concluded that the graphical method is useful for studying the transient characteristics of a reactor, but, owing to the importance and complexity of the subject, experimental confirmation of the predicted results is desirable.

LIST OF PRINCIPAL SYMBOLS

Neutron Kinetics.

k_{eff} = Effective reproduction constant.

$k_{ex} = k_{eff} - 1$.

k_{exa} = Value of k_{ex} applied at beginning of a transient.

l = Life-time of the prompt neutrons, sec.

n = Neutron concentration at a given point within the reactor.

r_i = Density of species i at the same point as n.

t = Time, sec.

β = Fraction of fissions which result in production of delayed neutrons.

λ_i = Relaxation time for latent nuclei of species i, sec^{-1}.

μ_i = Fraction of the fissions producing delayed neutrons which result in nuclei of species i.

ω_1 = Reciprocal of the reactor steady-state period for the condition k_{ex} = constant, sec^{-1}.

ω_2 = Addition necessary to ω_1 to obtain the reciprocal of the reactor period under conditions when $\dot{k}_{ex}$ = constant, sec^{-1}.

Thermodynamic and Design Parameters.

C = Specific heat, kW-sec/deg C-lb.

C_1, C_2, C_3 = Thermal capacities of components, kW-sec/ft-deg C.

$$F = \psi W C_c + \frac{(L/2 + z)}{R_2}$$

h_1 = Heat-transfer coefficient, fuel to gas based on the temperature drop $(T_{us} - T_c)$, kW/deg C-ft^2.

h_2 = Heat-transfer coefficient gas to graphite, kW/deg C-ft^2.

K = Channel thermal characteristic, deg C/kW.

L = Length of channel, ft.

L_1 = Extrapolated length of channel, ft.

P = Total power developed in a channel, kW.

P' = Rate of heat production per unit length, kW.

p_1 = Effective periphery for heat transfer of can, ft.

p_2 = Effective periphery for heat transfer at graphite wall, ft.

r = Radial distance of a channel from core centre, ft.

R_1, R_2 = Resistances to heat flow, deg C-ft/kW.

S = Effective cross-sectioned area (Fig. 1), ft^2.

T = Temperature above the coolant inlet temperature at a point z along the channel (measured from the centre of the channel), deg C.

$\bar{T}$ = Temperature above the coolant inlet temperature, up to the point in the channel considered, defined by eqn. (3), deg C.

T_1 = Temperature defined by eqn. (16).

u = Coolant velocity, ft/sec.

W = Mass flow rate of coolant gas in channel, lb/sec.

ρ = Density, lb/ft^3.

τ = Channel fuel-element time-constant, sec.

ψ = Ratio of coolant temperature at point z to mean value of coolant temperature up to the point.

Suffixes.

The suffixes u, c and g indicate quantities applying to the fuel rod, the coolant and the graphite moderator, respectively.

The suffix 0 indicates initial conditions.

The suffix s indicates value at the surface.

(1) INTRODUCTION

The gas-cooled reactor will be used for the first stage of the British nuclear-power programme in generating stations of which Calder Hall is the prototype. In the design and performance of these plants, the temperature of the fuel-element sheath is a critical factor. For reasons of thermodynamic performance, this should be as high as possible in normal running; but in setting the limit, both normal operation and temporary periods of fault require consideration. With present designs and usages of material, the general consequences of exceeding the limit would be mechanical failure of the fuel-element sheath accompanied by release of fission products to the coolant gas, followed by the relatively slow oxidation of the uranium. The fission products would be contained within the gas-tight pressure envelope, but the internal contamination of the reactor might be severe. Such a possibility requires critical examination, taking account, so far as possible, of the inherent characteristics of the reactor, as opposed to the superimposed devices for control and tripping.

The paper briefly describes a number of transient conditions which are likely to be important in gas-cooled reactors, and

Mr. Bowen is in the Industrial Group of the United Kingdom Atomic Energy Authority.

Mr. Masters is with Preece, Cardew and Rider, and was attached to the Industrial Group of the U.K.A.E.A. when the paper was prepared.

suggests methods of calculating the consequent variation of reactor temperatures. The situation may be complicated by associated events peculiar to a given reactor design, such as fire, release of stored crystal energy or deterioration of heat-transfer surfaces. These possibilities require to be taken into account in the interpretation of the results yielded by the simple theory discussed.

(2) CONDITIONS CAPABLE OF CAUSING TEMPERATURE TRANSIENTS

In equilibrium, the rates of heat production and heat removal are balanced. A change from equilibrium due to either an increase of heating power or a reduction in cooling power produces a condition that is potentially dangerous.

(2.1) Increased Heating Power

The power produced by a nuclear reactor depends upon the reactivity as a function of time. One method of initiating a change of reactivity is by moving the control rods, but in the present context additional possibilities should be considered, e.g. structural failure or the inadvertent withdrawal of some absorber other than a control rod, leading to a reactivity release. The resulting changes in temperature and power modify the reactivity, owing to temperature coefficients and xenon poisoning respectively, and so a train of events ensues. Changes of reactivity due to temperature are examined in the following Sections, but changes of xenon poisoning, being slow compared with temperature effects, have been neglected.

(2.2) Decrease of Cooling Power

Heat is transferred from the reactor to water and thence removed as steam by the circulation of an intermediate heat-transfer gas, generally carbon dioxide. For a given gas, the heat transfer depends on the mass flow rate of the gas, which is determined by the circulator speed and the gas pressure. Cases are examined in the paper arising from assumed failure of the circulating machinery and loss of pressure. In a practical system, any such event would be arranged to initiate a controlled reactor shut-down by suitable movement of the control rods; in the paper the assumption is made that, for unspecified reasons, the control rods do not move. This is normally an unjustifiable assumption, and is made only to permit the analysis of the intrinsic safety of the reactor.

(2.3) Thermal Shock

The problem of the fall in temperature occasioned by an emergency reactor shut-down may be treated by methods similar to those developed below. Hence, the stresses in the fuel elements, graphite moderator and reactor structure due to the thermal shock may be estimated.

(3) FACTORS AFFECTING TRANSIENT PERFORMANCE

If the effective multiplication factor differs from unity, the neutron population, and hence the power, varies with time. In principle, the power change may be calculated from the equations which are often referred to as the 'neutron kinetic equations'. Apart from k_{eff}, the data required in the neutron kinetic equations are the neutron life-time and the delayed-neutron abundances.[1,2] The former has a characteristic value according to the type of reactor, i.e. it depends mainly on whether a fast reactor, a graphite-moderated, a heavy-water- or a light-water-moderated thermal reactor is considered. The delayed-neutron abundances differ according to the reactor fuel, depending on whether uranium 235, plutonium 239 or uranium 233 is used. Thus all reactors of the same class, e.g. graphite-moderated natural-

uranium reactors, have the same kinetic behaviour following a similar change of reactivity.

However, the resulting reactivity after the transient has commenced is not the same as the impressed value. For example, suppose the transient to be examined is due to a sudden movement of the control rods, such that k_{eff} increases from unity to $(1 + k_{exa})$, where k_{ex} is the excess multiplication factor, and the suffix a denotes that it is the original, or applied, value. At any subsequent time, k_{ex} equals the sum of k_{exa} and a contribution due to the consequent changes occurring within the reactor. Among these, only temperature effects significantly influence the short-term transient behaviour.

An explanation of the variation with temperature of the reactivity of a reactor is given in text-books on reactor physics.[3] In thermal reactors the effects may be divided into those associated with the temperatures of the fuel and the moderator respectively. It is unlikely in the practical case that either of these temperatures will be uniform throughout the system, but it is possible, under suitable conditions which are discussed further in Section 9.2 et seq., to choose a representative value for each component.

Once the effect of reactivity on power and the effect of temperatures on reactivity have been considered, the final link in the sequence is the effect of power on temperatures, which may be studied by considering an elementary cell of a typical thermal reactor, such as that illustrated in Fig. 1. The transient tem-
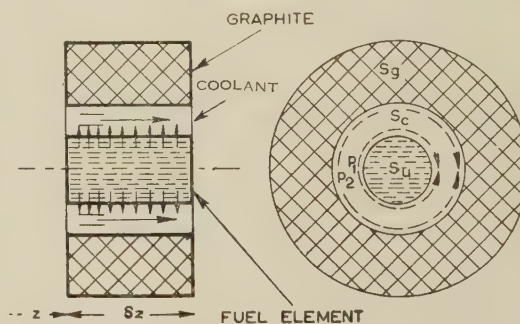
Fig. 1.—Simple model, unit cell of uranium, coolant and graphite.

perature changes depend on the thermal capacities of the components, the processes of heat exhange between the components and the transient variation of power and coolant flow. In contrast to the neutron-kinetic equations, only very limited generalization is possible for the thermal process, which is closely conditioned by the design of the reactor.

(4) METHODS OF SOLUTION

The interaction of so many effects presents a difficult problem in analysis, made more formidable by the form of the neutron-kinetic differential equations. Machines have been devised[8] in which the neutron kinetic equations are simulated in analogue form, but engineering feasibility sets a limit to the accuracy of such a device and requires the introduction of simplifying assumptions. Perhaps the most important arises from the fact that, in contrast to the variation of neutron density which occurs effectively simultaneously throughout the reactor, variation of temperature spreads in wave fashion with a transit time which may be significant. A practical analogue, however, represents conditions at one point, or at best at a few points, within the reactor.

If a digital computer is used to solve the problem, the only limits to accuracy are the effort required in programming the calculations and the machine time in performing them. However,

eat accuracy would be misplaced in the present context, both ecause the original fault conditions cannot be precisely defined, nd because much of the data (e.g. reactivity temperature and eat-transfer coefficients) are known only within a few per cent. ne need is for a reproducible method of analysis which gives onsistent results when applied over a wide range of values of e fundamental parameters, so that the safety of different cases ay be assessed on a comparable basis and consistent margins ay be allowed. Thus, the potentially high accuracy of the igital method is largely offset, and its only justification is to low a reasonably complete representation of the distribution space.
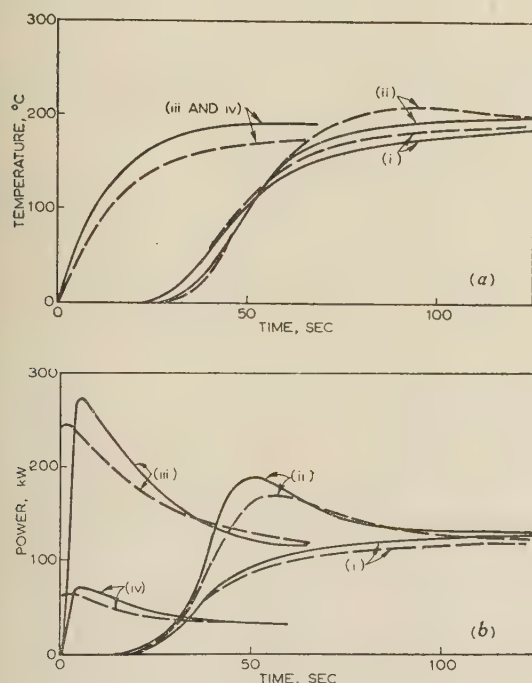
g. 2.—Power and temperature transients, 0·4% change in reactivity.
   (i) $K = 1·6 \deg C/kW$, $P_0 = 0·02 kW$, $\tau = 8 sec$.
   (ii) $K = 1·6 \deg C/kW$, $P_0 = 0·02 kW$, $\tau = 24 sec$.
   (iii) $K = 1·6 \deg C/kW$, $P_0 = 200 kW$, $\tau = 8 sec$.
   (iv) $K = 6·8 \deg C/kW$, $P_0 = 50 kW$, $\tau = 8 sec$.
   ——— Analogue results.
   - - - - Graphical results.
   (a) Temperature.
   (b) Power.
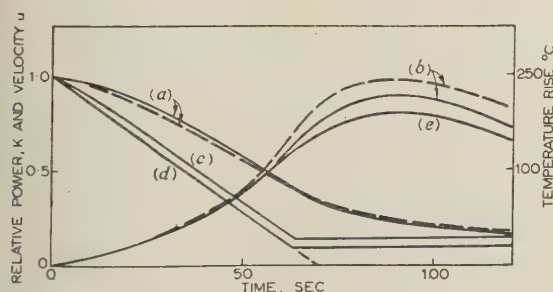
3.—Power and temperature transients following a steady fall in coolant velocity to 10%: pressure constant.
   ——— Analogue results.
   - - - - Graphical results.
(a) Power.
(b) Temperature.
(c) K_0/K.
(d) u/u_0.
(e) Temperature when pressure falls to 10% with coolant velocity constant.

If the equations representing the processes described in Section 3 are suitably modified by simplifying assumptions, graphical and analytic solutions may be developed. Such methods are described in Section 9 and the results obtained for various cases of practical importance are compared with analogue-computer solutions in Figs. 2 and 3 and in Table 1.

Table 1

Cause of change in cooling		Rate of change of cooling	Maximum transient mid-channel uranium temperature (starting temperature 258°C)	
Change in circulator speed	Change in coolant pressure		Analogue method	Graphical method
%	%		degC	degC
90	0	Step	470	502
90*	0	Step	388	
90	0	T.C. = 24 sec	429	425
90*	0	T.C. = 24 sec	347	
90	0	T.C. = 48 sec	380	
90	0	4·2%/sec	462	
90	0	1·43%/sec	425	440
0	90	Step	470	
0	90	4·2%/sec	436	
0	90	1·43%/sec	418	440
0	90	1·05%/sec	395	
80	0	Step	400	
80	0	T.C. = 24 sec	377	
70	0	Step	260	
70	0	T.C. = 24 sec	346	
50	0	Step	315	320
50	0	T.C. = 24 sec	307	
30	0	Step	286	294
30	0	T.C. = 24 sec	384	

* For plutonium fuel.

Note.—Rates of change of cooling in terms of time-constant (T.C.) refer to exponential run-downs.

(5) DISCUSSION OF RESULTS

The transient characteristics of the Harwell reactor Bepo have been discussed by Moore.[1] With the methods given in Section 4, similar characteristics have been calculated for a range of designs of graphite-moderated gas-cooled reactors, and from this study some general conclusions have been formulated on the effects of different variables on three aspects of transient behaviour which are important in practice.

(5.1) Abrupt Increase of Applied Reactivity, with Constant Cooling Conditions

This may be discussed in terms of the power, P, produced by a given channel in a particular reactor, and a mean temperature, $\bar{T}_u$, which is representative of the fuel elements. Cooling conditions may be specified by a parameter, K, which in equilibrium equals $\bar{T}_u/P$. Discussion of the transients is facilitated by a further parameter, which will be referred to as the channel time-constant, τ.

The typical course of events is shown in Fig. 2. As a consequence of the reactivity release, the power rises, and the uranium temperature, which follows, ultimately stabilizes at a value determined by the magnitude of the reactivity step and the temperature coefficient of reactivity. By definition, the change in equilibrium power is then $1/K$ times the final change in uranium temperature. However, before equilibrium is reached, power and temperature may overshoot or even oscillate. The influence of the various factors on the transient behaviour is illustrated in the four cases in Fig. 2, each of which represents a reactivity

increment of 0·4%, and the results may be summarized as follows.

The tendency to overshoot becomes greater if τ is increased, but a change in K does not alter the temperature overshoot. For given values of K and τ the transient is more severe the greater the power at which it commences.

For a given reactor channel, τ and K are not independent, as may be seen from eqn. (15), in which τ/K depends only on the thermal capacity of the fuel rods in the channel. For this reason, a power reactor working on part load, when it would be operating at a higher value of K and therefore a higher value of τ, has a greater tendency to overshoot. This is the explanation for the oscillatory response of Bepo noted by Moore (Fig. 4) as com-

Fig. 5.—Typical start-up curves for constant-speed rod withdrawal.
(*a*) Starting power of 1 watt.
(*b*) Starting power of 1 mW.

start-up, and the unknown critical point merely has the effect of altering the time at which the transition illustrated in Fig. 5 is enacted.

The start-up curves plotted in Fig. 6A show how the power varies as a function of the reactivity released at any time for a

Fig. 4.—Power and temperature transients of Bepo, following 0·2% k_{ex}, with low starting power.

pared with the cases illustrated in Fig. 2. The fuel-rod dimensions are very similar in both cases, but the value of K for Bepo is higher (Bepo: $P = 6$ kW, $\bar{T}_u = 300°$C, $K = 50$. Typical power reactor on full load: $P = 200$ kW, $\bar{T}_u = 200°$C, $K = 1$).

(5.2) Reactivity Increasing at a Constant Rate

This case is important in connection with starting up the reactor from its shut-down state, when the power is raised from the quiescent to the operating value. If the control-rod position where the reactor would be exactly critical could be predetermined, the start-up process could be halted at or near this point; but this point depends among other things on the previous operating history of the reactor and is not easy to calculate. If the control-rod withdrawal is continued at constant speed past the critical point, which in practice will not be evident, for a given control-rod speed (which approximates to a given rate of increase of reactivity) the rise of power will always follow a similar form, such as that shown in Fig. 5. During the first phase, until the reactor is slightly supercritical, power rises very slowly and is not visible on the scale to which Fig. 5 is drawn. In the second phase, power rises at an increasingly rapid rate. In the third phase, the rise of temperature stabilizes the situation and the final rate of rise of temperature and power is limited by the temperature coefficients of reactivity. Thus, by limiting the control-rod speed it is possible to ensure a definite pattern of

Fig. 6A.—Start-up curves without negative temperature coefficient effects.

——— Uranium 235 fuel.
- - - - Plutonium 239 fuel.

range of rates of reactivity increase. These curves, which were obtained with the use of the neutron-kinetic section of the analogue computer, are independent of the thermal characteristics, and therefore apply to all reactors of the same type (i.e. having the same mean neutron life-time and using similar fuel, which in this case is uranium 235). The broken curves in Fig. 6A represent the hypothetical case of a similar reactor using plutonium 239 as the fuel. In practical power reactors a part of the initial charge of natural uranium tends to be converted during operation to plutonium 239, so that the characteristics of a mature reactor may lie between the two cases.

After temperatures have risen significantly, the subsequent behaviour of the two systems is shown in Figs. 6B and 6C, in which the abscissa is reactivity released at any time, while the ordinate represents the reactivity taken up by temperature effects at that time. At a given abscissa the difference between a curve and the line ON represents the available reactivity which, for fast reactivity rates, may be quite large, towards the right-hand side of the Figure.

The conditions which determine whether the temperature over-
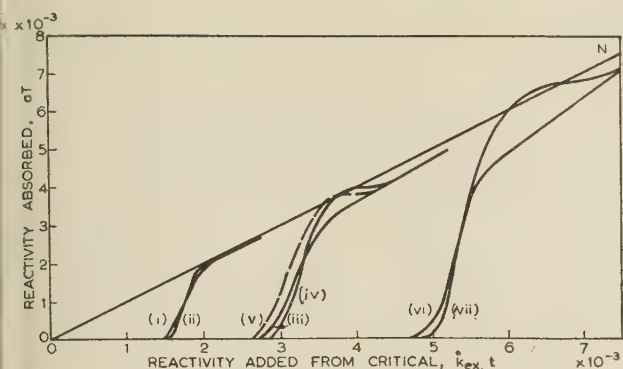
Fig. 6B.—Start-up curves for uranium 235, including negative temperature coefficient effects.

Curve	$\dot{k}_{ex}$	K	τ
	$\times 10^{-5}$	deg C/kW	sec
(i)	0·2	1	8
(ii)	0·2	1	40
(iii)	1	1	8
(iv)	1	1	40
(v)	1	10	40
(vi)	5	1	8
(vii)	5	1	24
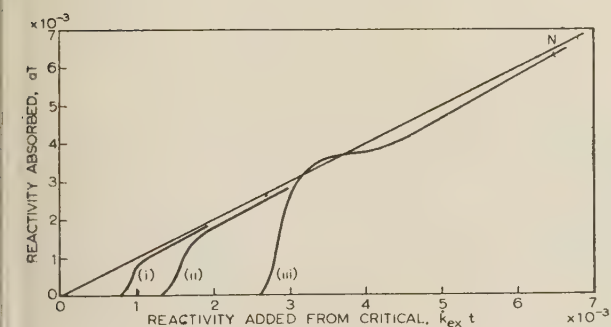
Fig. 6C.—Start-up curves for plutonium 239, including negative temperature coefficient effects.

(i) $\dot{k}_{ex} = 0·2 \times 10^{-5} k_{ex}/\text{sec}.$
(ii) $\dot{k}_{ex} = 1 \times 10^{-5} k_{ex}/\text{sec}.$
(iii) $\dot{k}_{ex} = 5 \times 10^{-5} k_{ex}/\text{sec}.$

oots before asymptotically approaching the line ON are, in neral, the same as those described in Section 5.1. The starting ower, denoted as P_0 in Fig. 6A, is always small (less than a few atts for the whole reactor), and overshoots are therefore unimortant over the range of practical values of the channel timeonstant. Although small, the starting power is capable of ing varied within limits in the design of the reactor, and quires to be taken into account. If, for example, starting ower in the case of a given curve in Fig. 6A were reduced by ne decade, then, since the power would have to rise by a further ecade before achieving the same effect on temperature, the rresponding curve in Fig. 6B would be displaced to the right this time interval. The result would be a proportionally eater surge of temperature. On the assumption that a limiting lue for the amplitude of temperature surge has been specified, ll be noted that, if the starting power were increased, e.g. by ng an artificial neutron source,[2] a faster start-up would be sible, since not only is the range over which power requires raised reduced, but also a higher reactivity rate is permisJudged by the criterion of temperature surge, it is seen m Fig. 6C that the change to plutonium fuel leads to easier

starting conditions, and it would appear that plutonium build-up in a power reactor need not be taken into account in determining control-rod speeds.

(5.3) Coolant Failure

After a nuclear reactor has been shut down, the continuing occurrence of radioactivity throughout the core, and in particular in the fuel elements, renders essential the provision of some residual cooling to remove the after-heat. Greatest reliability would be provided by natural circulation; where this is not technically feasible, either storage-battery or steam-driven back-up circulators may be employed.

As a consequence of coolant failure whilst at power, fuel-element temperature rises, reactivity is in turn reduced and the power subsequently falls. This continues until the cooling process stabilizes, after which power also stabilizes at a value such that temperatures are restored approximately to their original value. This is possible only if the residual cooling capacity is greater than is required by the after-heat at the original temperature levels.

These events are illustrated in Fig. 7. The effect of the para-
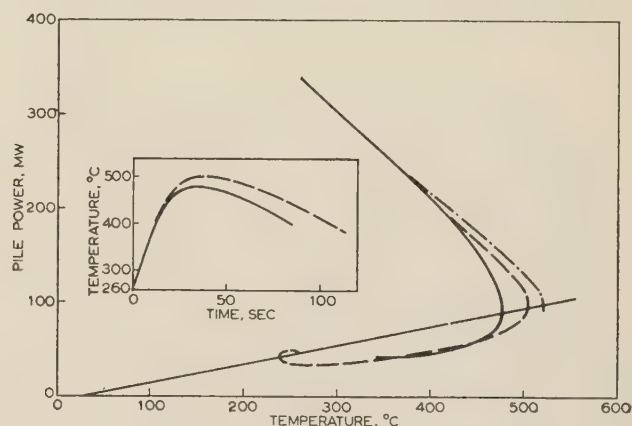
Fig. 7.—Power and temperature transients following step reduction of coolant velocity to one-tenth of rated value.

---- Graphical results.
—·— Graphical results including effect of fission-product heating.
——— Analogue results.

meters K and τ on the course of the transient is examined in more detail in Section 9, and it is shown that the severity of the transient depends on the rate at which the circulation fails. This is a function of the inertia of the rotating parts, while in the special case of steam-driven circulators it may also be retarded by the provision of a steam accumulator. By controlling such factors, the designer may limit the amplitude of the temperature transient.

If the cooling potential were reduced as a consequence of gas leakage from the pressure circuit, the range over which this could occur is given by the initial pressure expressed in atmospheres. An example of this type of transient is illustrated in Fig. 3(e). The result may suggest the advisability, for reasons of safety, of limiting the maximum operating pressure.

(6) CONCLUSIONS

Cases have been examined which might arise on gas-cooled reactor plants in consequence of severe fault conditions. Methods of calculation have been suggested which may be applied in such cases, provided that a number of simplifications are made. The results are not so accurate as those obtainable from the best methods of mechanical computation, but it is suggested that the

methods are useful in yielding reproducible results which correctly indicate the trends due to variation of the parameters. The use in the design stages of such methods for examining the transient characteristics of a projected plant should enable performance to be optimized without jeopardizing inherent safety under conditions of major faults.

Many of the cases examined in the paper are illustrated by analogue-computer solutions, and the calculations agree with these satisfactorily. Very few results are available from actual reactors, since the occurrence of such faults is extremely rare. Nevertheless, the results are not without importance in establishing confidence in the safety of these plants; and having in mind the complexity of the subject, it would appear justifiable to seek further confirmation of the predicted results by experiments on actual reactors.

(7) ACKNOWLEDGMENTS

The authors thank the Managing Director of the Industrial Group, United Kingdom Atomic Energy Authority, for permission to publish the paper. They would also like to acknowledge the help given by Mr. I. Wilson and Mr. E. Childs with the part of the investigation carried out on the Harwell analogue computer.

(8) REFERENCES

(1) MOORE, R. V.: 'The Control of a Thermal-Neutron Reactor', *Proceedings I.E.E.*, Paper No. 1431 M, December, 1952 (**100**, Part I, p. 90).

(2) COX, R. J., and WALKER, J.: 'The Control of Nuclear Reactors', *ibid.*, Paper No. 2068 M, March, 1956 (**103** B, p. 577).

(3) GLASSTONE, S., and EDLUND, M. C.: 'The Elements of Nuclear Reactor Theory' (Macmillan, London, 1953).

(4) HURWITZ, H.: 'Derivation and Integration of the Pile Kinetic Equations', *Nucleonics*, 1949, **5**, p. 61.

(5) WOODROW, J.: 'Thermal Time Lag in Air-Cooled Pile', A.E.R.E. Report No. E/R 142.

(6) GILLESPIE, A. B.: 'The Control and Instrumentation of a Nuclear Reactor', *Proceedings I.E.E.*, Paper No. 2058 M, March, 1956 (**103** B, p. 564).

(7) BOWEN, J. H.: 'Automatic Control Characteristics of Thermal-Neutron Reactors', *ibid.*, Paper No. 1432 M, December, 1952 (**100**, Part I, p. 102).

(8) MACLUSKY, G. J. R.: 'An Analogue Computer for Nuclear Power Studies', *ibid.* Paper No. 2337 M, March, 1957 (**104** B, p. 433).

(9) APPENDICES

(9.1) Development of Graphical Solutions

(9.1.1) The Reactor Kinetic Equations.

An explanation of these equations is given in Reference 1 and a rigorous derivation in Reference 4. They are:

$$dn/dt = [k_{ex}(1 - \beta)n]/l + \sum_{i=1}^{i=6} \lambda_i r_i \quad . \quad . \quad . \quad (1)$$

$$dr_i/dt = k_{ex}\mu_i\beta n/l - \lambda_i r_i \quad . \quad . \quad . \quad . \quad (2)$$

$$\sum_{i=1}^{i=6} \mu_i = 1$$

(Note: No concise term appears to have been coined for k_{ex} = reproduction constant minus one. Reactivity is strictly defined as k_{ex}/k_{eff}; however, in the class of reactor described here, k_{eff} is so nearly unity that the term reactivity may be applied to k_{ex}.)

A correction is necessary to the value of power calculated from eqns. (1) and (2), owing to the phenomenon of after-heat. This arises from the continuing radioactive decay of the products of past fissions, and for the time intervals considered here, may be taken as a constant, equal to about 5% of the steady power which applies at the commencement of the transient conditions. This constant component has been ignored in all the subsequent Sections except 9.1.6, where a method of applying a correction is given.

The values of the fission yields, $\mu_i\beta$, of neutron-emitting isotopes vary according to the parent fuel, and the kinetic behaviour is correspondingly different. Values for uranium 235 and plutonium 239 are given in Table 1 of Reference 2.

(9.1.2) Interaction of Temperature and Reactivity.

The phenomena which cause these effects are described in Reference 3. The results may be summarized for present purposes by assuming that the reactivity absorbed is proportional to the sum of the products of the temperature and partial temperature coefficients for the graphite and fuel separately, and for a hypothetical reactor in which the neutron flux and the temperature of a given component are uniform throughout the system,

$$k_{ex} = -\alpha_u T_u - \alpha_g T_g$$

In practice, the axial distribution of temperature and to a smaller extent the radial distribution are not uniform, and therefore the net effect of these different temperatures on k_{ex} is required. It is shown in Reference 3 that the effect of temperature at a given point should be weighted in proportion to the square of the neutron density at that point.

In addition to the variation in temperature along the channels and across the reactor, a radial temperature gradient exists within the uranium bar. Calculation of the effect on reactivity of temperature within the bar is complicated by the fact that both neutron density and neutron energy change with depth of penetration into the bar. This effect is believed to be insignificant in the present context, and it is assumed that the reactivity change is proportional to the change in mean temperature over the bar cross-section at any point along its length. If T_{us} is the surface temperature at this point, and $T_{us} + T$ is the temperature of an element of the bar cross-section S at this point, the internal mean temperature T_{um} is defined by

$$T_{um} = T_{us} + \frac{1}{S}\int^{S} T dS$$

In the following Sections an axial mean temperature $\bar{T}_u$ is employed, defined by:

$$\bar{T}_u = \frac{1}{(L/2 + z)}\int_{-L/2}^{z} T_{um} dz \quad . \quad . \quad . \quad . \quad (3)$$

Therefore, if the form of the distribution of temperature and neutron density remains unchanged during a transient, k_{ex} may be related to any convenient temperature, e.g. to $\bar{T}_u$,

$$k_{ex} = k_{exa} - a_u(\bar{T}_u - \bar{T}_{u0}) - a_g(\bar{T}_g - \bar{T}_{g0}) \quad . \quad . \quad (4)$$

a_u bears a constant relation to α_u given by

$$\frac{a_u}{\alpha_u} = \frac{\int T_{um} n^2 dv}{\bar{T}_u \int n^2 dv} \quad . \quad . \quad . \quad . \quad (4a)$$

Similar expressions apply for a_g.

The condition that the neutron density distribution remains constant is usually satisfied, since it is little affected by the reactivity redistribution which can occur during a practical transient. The condition that the redistribution of temperature

ring a transient does not affect a_u requires checking for a
ven case.

1.3) Heat-Balance Equations.

The process of heat exchange between the fuel rods, the coolant
s and the graphite matrix may be studied using the simplified
odel shown in Fig. 1. An elemental length of one channel
considered, in which the circular boundary of the graphite
ll is chosen to give the correct volume of graphite associated
th the element. Zero flow of heat is assumed at the boundary
the cell, and diffusion time into the graphite is neglected, i.e.
iform temperatures are assumed over any plane perpendicular
the channel axis. The design of the type of reactor dealt
th here is such that these conditions are approximately satisfied
equilibrium, and under transient conditions are satisfied within
e limits discussed in Section 9.4. Consider the heat balance
r each of the elements of the cell shown in Fig. 1.

ranium. $C_1 \dfrac{\partial T_{um}}{\partial t} = \hat{P}' \cos \pi z/L_1 + (T_c - T_{um})/R_1$ . . (5)

oolant.

$$C_2(\partial T_c/\partial t + u\partial T_c/\partial z) = (T_{um} - T_c)/R_1 + (T_g - T_c)/R_2 \quad (6)$$

oderator. $C_3 \partial T_g/\partial t = (T_c - T_g)/R_2$ . . . . (7)

here

$$C_1 = \rho_u C_u S_u$$
$$C_2 = \rho_c C_c S_c$$
$$C_3 = \rho_g C_g S_g$$
$$R_1 = 1/p_1 h_1 + (T_{um} - T_{uS})/P'$$
$$R_2 = 1/p_2 h_2$$

These equations were given by Woodrow,[5] who justified the
glect of conduction in the axial direction compared with con-
ction heat transfer. The temperatures refer to the difference
tween the values measured at a given point and that of the
olant gas at the inlet to the reactor. The inlet temperature is
sumed to be constant throughout the transient, which is
stified by the approximately constant temperature charac-
ristics of the heat exchangers.

In practice, the fuel material is surrounded by a sheath having
thermal capacity and a thermal resistance between it and the
el. In practical designs the latter is small compared with R_1
d may be included with it, the sheath thermal capacity being
mped with that of the fuel. The thermal time-constant of the
eath is neglected by this procedure.

1.4) Simplifying Approximations.

Eqns. (5), (6) and (7) are partial differential equations, but
ese may be converted to ordinary differential equations by
tegrating both sides of each with respect to z. In effect, T_{um},
and T_g are then replaced by their mean values $\bar{T}_u$, $\bar{T}_c$ and $\bar{T}_g$
ong the channel up to the point z considered. This is per-
issible since the other variables in eqns. (5), (6) and (7), namely
R_1, and R_2, do not vary along a given channel, although this
not necessarily true between different channels (see Section 9.2).
e equations for the variation of mean temperatures are given
Section 9.3, where the example is solved of a step in power.
typical practical solution is plotted in Fig. 9. In practice, it
necessary to reinterpret the mean temperatures in terms of
tual values in a channel. To a first approximation it may be
sumed that the temperature changes occur in phase at all
ints along the channel, when the steady-state distribution, such
shown in Fig. 8, may be used to determine the ratio between
e maximum temperature and the mean. The error introduced
this assumption is examined in Section 9.4.
s can be seen from Fig. 9, the moderator temperature does
t vary significantly during uranium temperature transients.
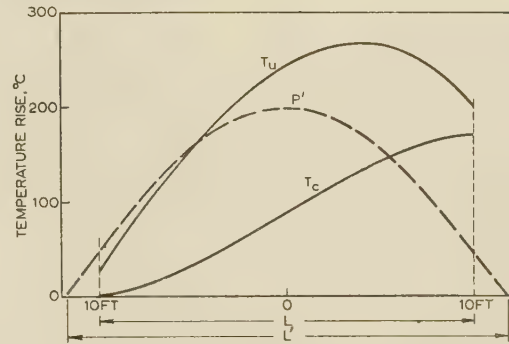
Fig. 8.—Distribution of power and temperature along a channel.
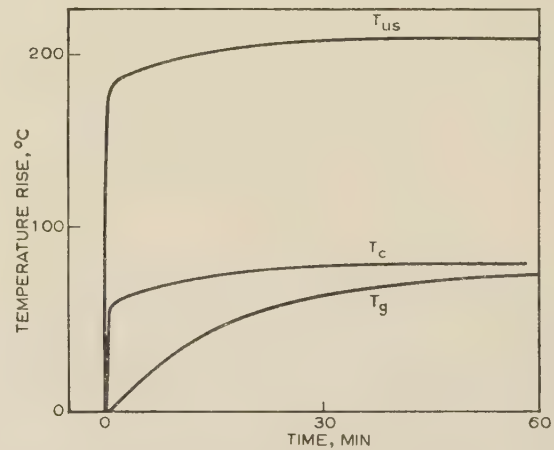
Fig. 9.—Variation with time of fuel, coolant and moderator
temperature at a point in a channel.

Making this assumption, the simplified equations may be written

$$\bar{P}' - C_1 d\bar{T}_u/dt = (\bar{T}_u - \bar{T}_c)/R_1 \quad . . . . . . (8)$$

$$= \frac{\psi W C_c \bar{T}_c}{(z + L/2)} + \frac{\bar{T}_c - \bar{T}_{c0}}{R_2}$$

$$= F\bar{T}_c/(z + L/2) - \bar{T}_{c0}/R_2 \quad . . . (9)$$

For the most accurate calculation of maximum uranium tem-
perature the point z should be chosen to coincide with maximum
temperature. Owing to the symmetry of the heat distribution,
however, simpler expressions are obtained if $z = L/2$, i.e. the
complete channel is considered. Whether this approximation is
permissible in a given case may be judged by the criterion of the
temperature distribution.

The reactor-kinetic formulae, eqns. (1) and (2), have been
solved for the case in which the reactivity increases abruptly
to a given value from zero. It has been shown by Gillespie[6]
that the solution may be represented approximately by

$$P' = A(k_{ex})P_0'\varepsilon^{\omega_1 t} \quad . . . . . (10)$$

where $A(k_{ex})$ is a multiplying factor which is a function of the
reactivity increment and ω_1 is also a function of k_{ex}. For very
slow changes of k_{ex} Hurwitz[4] has developed an expression
similar to eqn. (10), but this condition does not apply to the
rates of change of reactivity which occur in temperature
transients. Bowen[7] has suggested an approximation derived from the fre-
quency characteristics of a reactor. Rewriting eqn. (10) of
Bowen's paper gives

$$\frac{1}{P'} \frac{dP'}{dt} + 0.2 \frac{d}{dt}\left(\frac{1}{P'} \frac{dP'}{dt}\right) = 30k_{ex} + 150\frac{dk_{ex}}{dt}$$

This equation has been employed in the present work with certain modifications. By suitably restricting the cases considered, it was found possible to ignore the term containing the second differential coefficient of the power; whilst the numerical coefficients of k_{ex} and $\dot{k}_{ex}$ have been replaced by functions of these variables. The function of k_{ex} which is referred to in eqn. (10) as ω_1 has been calculated for graphite-moderated thermal reactors and is shown in Fig. 10. The function of $\dot{k}_{ex}$

Fig. 10.—ω_1 as a function of k_{ex}.

(a) ω_1 and k_{ex} positive.
(b) ω_1 and k_{ex} negative.

has been obtained from experiments in which different rates of reactivity variation were applied to the analogue computer, and is shown in Fig. 11. The relationship is approximately

$$\omega_2 = 170\dot{k}_{ex}$$

and this may be compared with the value given above, which was obtained by a completely different method. The approxima-

Fig. 11.—ω_2 as a function of $\dot{k}_{ex}$.

tion employed in the calculations of the paper is in terms of $\bar{P}'$ given by

$$(1/\bar{P}')(d\bar{P}'/dt) = \omega_1(k_{ex}) + \omega_2(\dot{k}_{ex}) \quad . \quad . \quad (11)$$

(9.1.5) **Graphical Solution of Simplified Equations.**

Substituting for $\bar{T}_c$ from eqn. (9) into eqn. (8) gives

$$\frac{d\bar{T}_u}{dt} = \frac{1}{C_1}\left\{ \bar{P}' - \frac{\bar{T}_u - \dfrac{(z + L/2)\bar{T}_{c0}}{R_2 F}}{R_1\left[1 + \dfrac{(z + L/2)}{R_1 F}\right]} \right\} \quad . \quad . \quad (12)$$

Dividing eqn. (11) by eqn. (12) gives

$$\frac{d\bar{P}'}{d\bar{T}_u} = C_1\left\{ \frac{(\omega_1 + \omega_2)}{1 - \left[\dfrac{1}{R_1 + \dfrac{(z + L/2)}{F}}\right]\left[\dfrac{\bar{T}_u - \dfrac{(z + L/2)\bar{T}_{c0}}{R_2 F}}{\bar{P}'}\right]} \right\} \quad (13)$$

It is generally more convenient to rewrite this equation in terms of the design-point conditions. In the following it is assumed that it is sufficiently accurate to work in terms of the complete channel length, i.e. $z = L/2$. Then $\psi = 2$, and also $\bar{P}' = P/L$ where P is the power output from the channel. Then

$$\frac{dP}{d\bar{T}_u} = \frac{\tau/K}{1 - \dfrac{1}{K}\left(\dfrac{\bar{T}_u - T_1}{P}\right)} \times (\omega_1 + \omega_2) \quad . \quad . \quad (14)$$

where

$$\frac{\tau}{K} = C_1 L \quad . \quad . \quad . \quad . \quad . \quad (15)$$

$$T_1 = \frac{\bar{T}_{c0}}{1 + \dfrac{(P/\bar{T}_c)_{eq.0}(W/W_0)}{Lh/R_2 h_0}} \quad . \quad . \quad . \quad (16)$$

$$K = \left(\frac{\bar{T}_u - \bar{T}_c}{P}\right)_{eq.0}\left[\frac{(h_0/h)m + 1}{m + 1}\right] + \frac{\left(\dfrac{\bar{T}_c}{P}\right)_{eq.0}\dfrac{W_0}{W}}{1 + \dfrac{L}{R_2}\dfrac{h}{h_0}\left(\dfrac{\bar{T}_c}{P}\right)_{eq.0}\dfrac{W_0}{W}}$$

$$\quad . \quad . \quad . \quad (17)$$

$$m = \left(\frac{\bar{T}_u - \bar{T}_c}{\bar{T}_u - \bar{T}_{us}}\right)_{eq.0}$$

The trajectory of a point on the phase-plane of P and $\bar{T}_u$ gives the variation of P as a function of $\bar{T}_u$. Such a trajectory may readily be plotted, since the right-hand side of eqn. (14) contains as variables only P, $\bar{T}_u$ and $d\bar{T}_u/dt$. The latter can be indicated on the phase-plane, by lines of constant values of $d\bar{T}_u/dt$ according to the equation

$$P = \frac{\tau}{K}\frac{d\bar{T}_u}{dt} + \frac{1}{K}(\bar{T}_u - T_1) \quad . \quad . \quad . \quad (18)$$

A number of worked examples are shown in Fig. 12 for a transient initiated by a step of reactivity. The initial conditions associated with eqn. (10) are represented by the instantaneous step from P_0 to AP_0 which occurs at constant value of $\bar{T}_u$. These results are compared with analogue-computer solutions in Fig. 2 and indicate satisfactory agreement.

A worked example is shown in Fig. 7 for the case of a sudden reduction in cooling power to one-tenth of its initial value. Fig. 13 shows the two cases of an increase and a decrease in

Fig. 12.—Variation of power with mean uranium temperature for transients due to 0.4% k_{ex} step increases in reactivity.

Curve	P_0	K	τ	Curve of Fig. 2
	kW	deg C/kW	sec	
(a)	0·02	1·6	8	(i)
(b)	200	1·6	8	(ii)
(d)	0·02	1·6	24	(iii)
(c)	50	6·4	8	(iv)
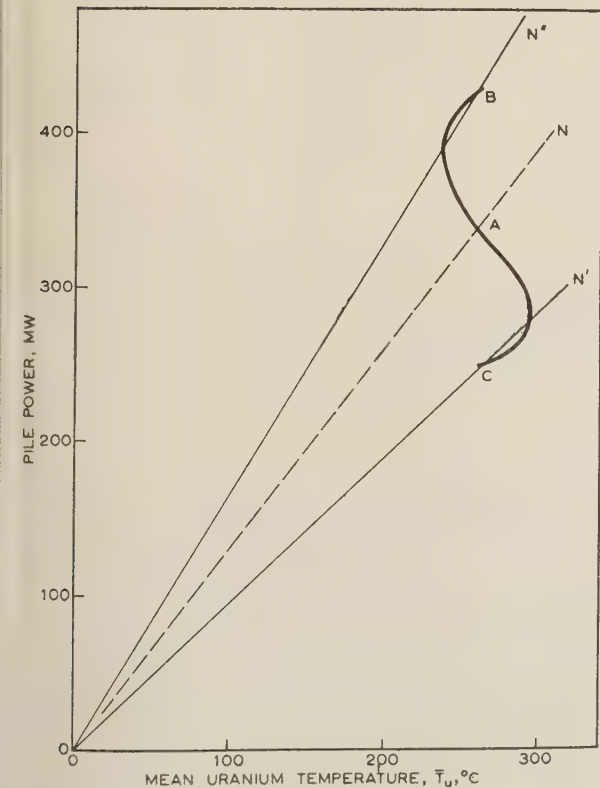
Fig. 13.—Variation of power with mean uranium temperature for transient due to 30% change in coolant flow.

A. Initial operation point.
B. Operating point after 30% increase in coolant flow.
C. Operating point after 30% decrease in coolant flow.

cooling power by 30%. Together the curves illustrate the principle of changing power isothermally by adjustment of cooling power, which is the normal method of changing the load on the Calder Hall reactors. The transient temperature changes may be made as small as desired by limiting the magnitude of the steps, or in practice by the gradual adjustment of cooling power. Some final trimming of reactivity is necessary if it is desired to maintain can surface temperature constant.

If it is desired to determine the transient temperatures in terms of can surface temperature, this can be accomplished by subtracting the temperature drop $(\bar{T}_u - \bar{T}_{us})$ from the transient effective temperature, the value of $(\bar{T}_u - \bar{T}_{us})$ being proportional to the power. This operation is shown as curve (b) in Fig. 17.

In practice, the cooling potential could not fall off abruptly, and the effects of a more gradual reduction are shown in Fig. 3. The calculated results were obtained by numerically integrating eqns. (11) and (18), and the results of a wide range of cases which were tested in which the reduction of flow occurred at different rates and followed different laws are summarized in Table 1. Good agreement was obtained between the calculations and the results from the analogue computer.

It may be shown to a first approximation that the speed, N, of a practical gas circulator follows the law

$$N(t) = \frac{\lambda}{\lambda + t} N_0$$

where λ is a constant related to the inertia, and t is the time following loss of driving power. The peak temperatures reached following power failure and circulator run-down to different minimum speeds have been calculated for a range of values of λ, and are indicated in Fig. 14. From studies of this type it is
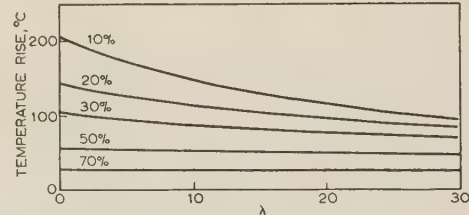
Fig. 14.—Transient temperature rises as a function of λ for run-down to minimum speeds of 10–70% of rated speed.

possible to estimate the circulator inertia which would satisfactorily limit the transient.

The previous solutions have been obtained in terms of the mean fuel temperature along the channel, but within the limits of the assumptions made in this Section the values of transient temperature changes occurring at the point of maximum temperature may be obtained by multiplying by the ratio of these quantities which holds in equilibrium. This is

$$\frac{\hat{T}_u}{\bar{T}_u} = \frac{\sqrt{\left[R_1^2 + \left(\dfrac{L_1}{\pi u C_2}\right)^2\right]} + \dfrac{L_1}{\pi u C_2}\sin\dfrac{\pi L}{2L_1}}{\dfrac{2R_1 L_1}{\pi L}\sin\dfrac{\pi L}{2L_1} + \dfrac{L_1}{\pi u C_2}\sin\dfrac{\pi L}{2L_1}} \quad . \quad (20)$$

(9.1.6) Effect of Fission-Product Heat on Transients.

As explained in Reference 2, a proportion of the operating power at any time is in the form of radioactive heating due to the γ and β decay of the fission products. The decay time of this activity is long compared with the time of the transient considered here, and it can therefore be assumed that, during the period of the transient of interest, a proportion of the power

remains constant. The result of this is to modify eqn. (11), which now becomes, in terms of channel power,

$$\dot{P}/(P - \Delta P_0) = \omega_1(k_{ex}) + \omega_2(\dot{k}_{ex}) \quad . \quad . \quad . \quad (21)$$

where ΔP_0 is the proportion of fission-product heat (assumed constant during the transient). Eqn. (14) is modified by a factor $1 - \Delta P_0/P$ and becomes

$$\frac{dP}{d\bar{T}_u} = \left[\frac{\tau/K}{1 - \frac{1}{K}\left(\frac{\bar{T}_u - T_1}{P}\right)} \right] \times (\omega_1 + \omega_2) \times \left(1 - \frac{\Delta P_0}{P}\right) \quad . \quad (22)$$

The effect of this modification in the case of a step fall of coolant down to $0 \cdot 1$ rated circulator speed is indicated in Fig. 7.

(9.2) Application to Practical Reactor Systems

The discussion in Section 9.1 applies to a particular channel in a reactor or to a hypothetical reactor in which all elements have similar heat outputs and temperatures. In practical reactors the heat output is not equal for all channels, but is usually least towards the edge of the core. The temperature depends on the heat output and on the coolant flow in a particular channel.

It would be possible to allow equal coolant flows through each channel irrespective of the heat release, e.g. in low-power experimental reactors in which neither economy of coolant pumping power nor development of high gas temperature was important. The temperature in each channel would be proportional to the respective heat outputs, and the value of the temperature/power ratio K would be uniform. Although of different amplitudes, the temperature transients would occur simultaneously in each channel, and if the temperature coefficient of reactivity were computed for $\bar{T}_u$ according to Section 9.1.2, the calculations already given would apply correctly to the central channel.

In power-generating reactors it is the practice to match the channel coolant flow to the heat release. The value of K is no longer uniform across the reactor, and if individual channels are treated separately according to the preceding discussion, the transients would appear to have varying forms and times depending on the value of K. In practice, the temperature transients probably do vary in time in different channels, being fastest when the power rating is highest. Temporarily, an abnormal distribution of reactivity prevails, but it has been explained in Section 9.1.2 that the distribution of power does not alter in consequence. The distribution of power and the variation of power with time are determined by the overall value of reactivity, averaged at any instant according to eqns. (4) and (4a). The power transient therefore tends to be determined by those channels which have most influence on the reactivity. If, as is usual, all the channels at a given radius from the core axis share the same conditions, the number of channels at a given radius, r, is proportional to $2\pi r$; and if the neutron concentration at that radius under normal equilibrium conditions is $n(r)$, the effectiveness of that group of channels is proportional to $2\pi r[n(r)]^2$. This quantity has a maximum value at a radius r^*, which may be determined for a given reactor (see Fig. 15), and the preceding analysis may be applied to a typical channel in the region r^*. The variation of power occurs simultaneously over the whole reactor, the amplitudes in the different channels corresponding to the equilibrium distribution of heat release. The temperature variation is given by the preceding calculation in the region r^*; in more highly rated channels, where K has a lower value, temperatures follow the power more closely; in the outer channels the lag is greater. The values may be calculated by solving eqns. (5), (6) and (7) for the temperature, the change in power as a function of time being given by the results derived
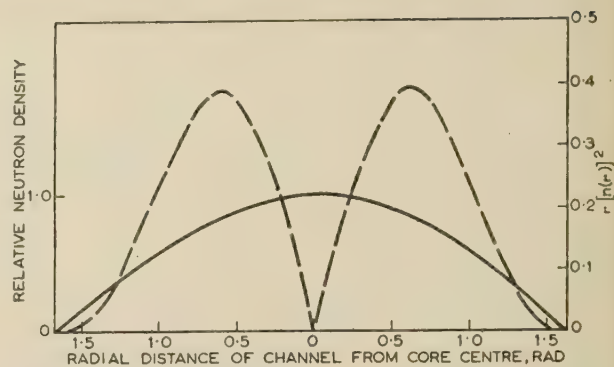
Fig. 15.—Variation of neutron density and channel effectiveness with radial distance from core centre.
——— Neutron density.
- - - - $r[n(r)]^2$.

from the region r^*; or, within the accuracy of that analysis, by similarly solving eqn. (18), the power once again being given. The steps in this procedure are illustrated by an example in Section 9.5.

(9.3) Solution of the Thermal Equations

An approximate solution may be obtained for eqns. (5)–(7) by noting that, in equilibrium, $\partial T_c/\partial z$ is not very different from the mean value of this quantity up to the point considered, i.e. T_c/z. If the criterion given in Section 9.4 is satisfied and temperature transients do not involve a transient redistribution of temperature, the mean value may be used in calculating the transients. Employing the following substitutions:

$$\theta = t/R_1 C_1$$
$$\gamma = C_1/C_2$$
$$\beta = C_1/C_3$$
$$\alpha = R_1/R_2$$
$$\nu = dz/d\theta$$

$$T = \frac{\hat{P}' R_1}{z + \frac{1}{2}L} \int_{-\frac{1}{2}L}^{+z} \cos\left(\frac{\pi z}{L_1}\right) dz$$

Eqns. (3)–(5) may be written

$$\frac{dT_u}{d\theta} + T_u - T_c = T \quad . \quad . \quad . \quad . \quad (23)$$

$$- \gamma T_u + \frac{dT_c}{d\theta} + T_c\left[\frac{\nu}{z} + \gamma(1 + \alpha)\right] - \alpha\gamma T_g = 0 \quad . \quad (24)$$

$$- \alpha\beta T_c + \frac{dT_g}{d\theta} + \alpha\beta T_g = 0 \quad . \quad . \quad . \quad (25)$$

Employing the Laplace transformation and denoting the variable of the Laplace transform by s gives

$$\tilde{T}_u = \left(\frac{T}{s}\right)\left\{s^2 + s\left[\frac{\nu}{z} + \gamma(1 + \alpha) + \alpha\beta\right] + \alpha\beta\left(\frac{\nu}{z} + \gamma\right)\right\}\Big/ D$$

$$\tilde{T}_c = \left(\frac{T}{s}\right)[\gamma(s + \alpha\beta)]/D$$

$$\tilde{T}_g = \left(\frac{T}{s}\right)(\alpha\beta\gamma)/D$$

here $\widetilde{T}$ indicates the Laplace transform of T, and

$$= s^3 + s^2\left[\frac{\nu}{z} + \gamma(1 + \alpha) + \alpha\beta + 1\right]$$

$$+ s\left[\frac{\nu}{z}(\alpha\beta + 1) + \alpha\beta(\gamma + 1) + \alpha\gamma\right] + \frac{\alpha\beta\nu}{z} \quad . \quad (26)$$

$$= (s - s_1)(s - s_2)(s - s_3)$$

here s_1, s_2, and s_3 are the roots of $D = 0$.

Then

$$= \sum_{n=1}^{n=3}\left\{s_n^2 + s_n\left[\frac{\nu}{z} + \gamma(1 + \alpha) + \alpha\beta\right] + \alpha\beta\left(\frac{\nu}{z} + \gamma\right)\right\}\frac{\varepsilon^{s_n\theta}}{D'}$$

$$+ \frac{\gamma z}{\nu} + 1 \quad . \quad (27)$$

$$\frac{T_c}{T} = \sum_{n=1}^{n=3}\left[\gamma(s + \alpha\beta)\frac{\varepsilon^{s_n\theta}}{D'}\right] + \frac{\gamma z}{\nu} \quad . \quad . \quad . \quad (28)$$

$$\frac{T_g}{T} = \sum_{n=1}^{n=3}\left(\alpha\beta\gamma\frac{\varepsilon^{s_n\theta}}{D'}\right) + \frac{\gamma z}{\nu} \quad . \quad . \quad . \quad (29)$$

here

$$' = s_n\left\{3s_n^2 + 2s_n\left[\frac{\nu}{z} + \gamma(1 + \alpha) + \alpha\beta + 1\right]\right.$$

$$\left. + \left[\frac{\nu}{z}(\alpha\beta + 1) + \alpha\beta(\gamma + 1) + \alpha\gamma\right]\right\}$$

(9.4) Effect of Channel Length on Transient Temperature Variation

Consider the simplified case of uniform heating along a channel with non-conducting walls. The equations of heat transfer corresponding to eqns. (5)–(7) become

$$C_1\frac{\partial T_u}{\partial t} = P' - \frac{(T_u - T_c)}{R_1} \quad . \quad . \quad . \quad (30)$$

$$C_2\left(\frac{\partial T_c}{\partial t} + u\frac{\partial T_c}{\partial z}\right) = \frac{(T_u - T_c)}{R_1} \quad . \quad . \quad . \quad (31)$$

These may be solved for the initial conditions, $t = 0$, $P' = 0$, and $T_u = T_c = 0$. Thereafter P' has a constant and uniform value. Putting $\theta = t/C_1R_1$, $\nu = \partial z/\partial\theta$, $\gamma = C_1/C_2$ and $T = R_1P'$, the equations become

$$\frac{\partial T_u}{\partial\theta} + T_u - T_c = T \quad . \quad . \quad . \quad . \quad (32)$$

$$\frac{\partial T_c}{\partial\theta} + \nu\frac{\partial T_c}{\partial z} - \gamma T_u - \gamma T_c = 0 \quad . \quad . \quad . \quad (33)$$

Eliminating T_u gives

$$\frac{\partial^2 T_c}{\partial\theta^2} + \frac{\partial}{\partial\theta}\left[\nu\frac{\partial T_c}{\partial z} + T_c(1 + \gamma)\right] + \nu\frac{\partial T_c}{\partial z} = \gamma T \quad . \quad (34)$$

Denoting the Laplace transform of T_c by $\widetilde{T}_c$ gives

$$s^2\widetilde{T}_c + s\left[\nu\frac{\partial\widetilde{T}_c}{\partial z} + \widetilde{T}_c(1 + \gamma)\right] + \nu\frac{\partial\widetilde{T}_c}{\partial z} = \frac{\gamma T}{s} \quad . \quad (35)$$

and solving for $\partial T_c/\partial z$ gives

$$\frac{\nu s}{\gamma T}\frac{\partial\widetilde{T}_c}{\partial z} = \frac{1}{s + 1}\exp\left[-\frac{z}{\nu}\frac{s(s + \gamma + 1)}{s + 1}\right]$$

which yields, on taking the inverse transformation

$$\frac{\partial^2 T_c}{\partial z\partial\theta} = \frac{\gamma T}{\nu}I_0\left\{2\left[\frac{\gamma z}{\nu}\left(\theta - \frac{z}{\nu}\right)\right]^{1/2}\right\}\exp-\left[\frac{\gamma z}{\nu} + \left(\theta - \frac{z}{\nu}\right)\right] \quad (36)$$

provided that $\qquad \theta > z/\nu$

The variation of $T_c/T_{c\infty}$ with θ, with $\gamma z/\nu$ as a parameter, is plotted in Fig. 16. The parameter $\gamma z/\nu$, which determines how
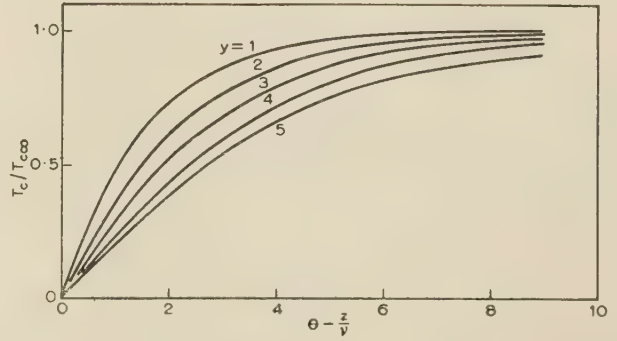
Fig. 16.—Coolant temperature rise plotted against normalized time, at points along the channel having transit times y expressed in multiples of $\gamma z/\nu$. As $y \to 0$ the form tends to exponential.

closely the behaviour agrees with the simple exponential form deduced in Section 9.3, may be given a physical interpretation: it is the time, expressed in multiples of the 'uranium thermal time-constant', taken by the coolant to reach the point z from the channel entry multiplied by the ratio of the thermal capacities per unit length of channel of uranium and coolant. Thus, at a point 20 ft along the reactor channel, with gas flowing at 40 ft/sec and a uranium thermal time-constant of 10 sec, $z/\nu = 0.05$. The value of γ varies from about 500 for an air-cooled reactor to 15 for a pressurized power reactor, so that the assumption made in Section 9.1 would be reasonably satisfied in the second, although not in the first, example.

(9.5) Example Illustrating Use of Formulae

(9.5.1) Determination of 'Most Effective Channel' Power and Temperature Transients.

The case to be examined is the addition of an 0.4% reactivity increase to a typical reactor when operating under rated cooling conditions and 10^{-3} of rated power.

Reactor specification:

Total pile power = 180 MW.
Total pile power − power in graphite = 93% of 180 = 167 MW.
Number of channels = 1 700.
Radial form factor = 1·89.
Therefore total central channel heat − heat to graphite = heat developed in fuel element = 186 kW.
Rated central channel mass flow = 2·54 lb/sec.
Rated central channel coolant velocity = 72 ft/sec.
p_2h_2 for central channel at rated conditions = 0·038 kW-sec/deg C-ft.
Radial neutron flux distribution assumed to be unflattened.

It is assumed that the equilibrium uranium and coolant temperatures are uniform in a plane perpendicular to the channel axis, and that the distribution along the channel axis is as shown in Fig. 8.

From Section 9.2 and Fig. 15 it can be seen that channels at a radial distance r^* from the core centre have the greatest weighted effect on transient power, where the channel power at r^* is 0.8 of the central channel value. Therefore the channel power at r^* is 150 kW.

It is assumed that the coolant mass flow is regulated across the

core in proportion to the variation in power; the mass flow in the channel under consideration will then be $2 \cdot 032$ lb/sec and the coolant velocity will be $57 \cdot 2$ ft/sec. The temperature coefficient, based on mean uranium temperature $\bar{T}_u$ for the most effective channel, is $-1 \cdot 5 \times 10^{-5} k_{ex}$/deg C. It is also assumed that the coolant pressure is constant at the rated value throughout the transient.

Starting Conditions.

$$P_0 = 0 \cdot 15 \, \text{kW, i.e. } 10^{-3} \times \text{rated power.}$$
$$\left.\begin{array}{l} \bar{T}_{u0} = 0 \cdot 245° \text{C} \\ \bar{T}_{c0} = 0 \cdot 086° \text{C} \end{array}\right\} \begin{array}{l} \text{scaled down from Fig. 8 in} \\ \text{proportion to the power.} \end{array}$$

The plot of power and temperature transients is carried out by a step-by-step method from the evaluation of $dP/d\bar{T}_u$ at consecutive points in the $P\bar{T}_u$ plane. To facilitate the tabulation of the step-by-step graphical construction it is convenient to express eqn. (14) in the form

$$dP/d\bar{T}_u = X(\omega_1 + \omega_2)$$

where

$$X = \frac{\tau/K}{1 - \frac{1}{K}\left(\dfrac{\bar{T}_u - T_1}{P}\right)} \qquad . \quad . \quad . \quad (37)$$

$$\omega_1 = \text{f}(k_{exa} - a_u\Delta\bar{T}_u)$$
$$\omega_2 = \text{f}(d\bar{T}_u/dt)$$

The value of ω_1 is obtained from Fig. 10, where the value of $\Delta\bar{T}_u$ is the temperature rise above the starting equilibrium condition at the instant considered. The value of ω_2 is obtained from Fig. 11, and the value of $d\bar{T}_u/dt$ is the rate of change of the mean fuel-element temperature along the channel at the instant considered. The value of $d\bar{T}_u/dt$ at any position of the plot is conveniently obtained from a system of lines of constant value of $d\bar{T}_u/dt$ shown on the power/temperature diagram according to the relation given in Section 9.1.5, which can be written as

$$\frac{d\bar{T}_u}{dt} = \frac{K}{\tau}P - \frac{1}{\tau}\bar{T}_u \quad . \quad . \quad . \quad . \quad (38)$$

Evaluating the parameters τ/K, K, τ, and T_1 gives

$$\tau/K = 7 \cdot 74 \, \text{kW-sec/deg C} \qquad K = 1 \cdot 48 \, \text{deg C/kW}$$
$$\tau = 11 \cdot 4 \, \text{sec} \qquad T_1 = 0 \cdot 0229° \text{C}$$

whence

$$X \simeq \frac{7 \cdot 74}{1 - 0 \cdot 675\bar{T}_u/P} \quad . \quad . \quad . \quad . \quad (39)$$

$$d\bar{T}/dt \simeq 0 \cdot 129P - 0 \cdot 088\bar{T}_u \quad . \quad . \quad . \quad (40)$$

By means of eqns. (39) and (40) the lines of constant X and $d\bar{T}_u/dt$ are drawn on Fig. 17 and the transient curve relating power to temperature is constructed by the step-by-step method. From Fig. 17 the time to reach consecutive points on the power/temperature diagram can be tabulated using the relation

$$\Delta t = \frac{\Delta\bar{T}_u}{d\bar{T}_u/dt} \quad . \quad . \quad . \quad . \quad . \quad (41)$$

The curves showing the variation of power and temperature with time are shown in Fig. 18.

(9.5.2) Determination of Central Channel Power and Temperature Transients.

For reasons given in Section 9.2, the power transient in the central channel is of the same form as that in channel r^* but increased in amplitude by $1 \cdot 25$. The power curve is plotted as curve (c) in Fig. 18.
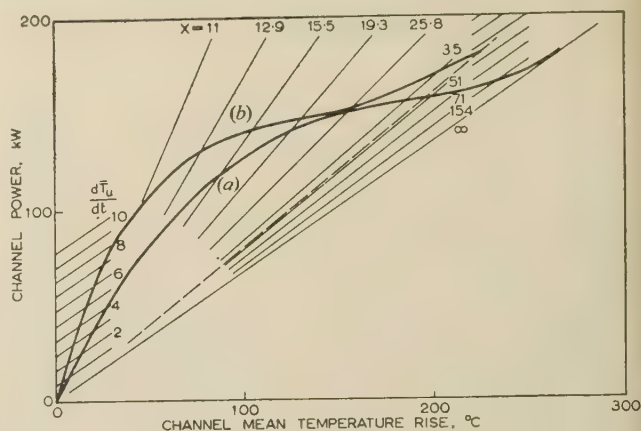
Fig. 17.—Variation of power with temperature for $0 \cdot 4\%$ k_{ex} step increase, showing construction lines of constant value of X and $d\bar{T}_u/dt$.

    (a) Transient based on channel mean effective uranium temperature.
    (b) Transient based on channel mean surface uranium temperature.
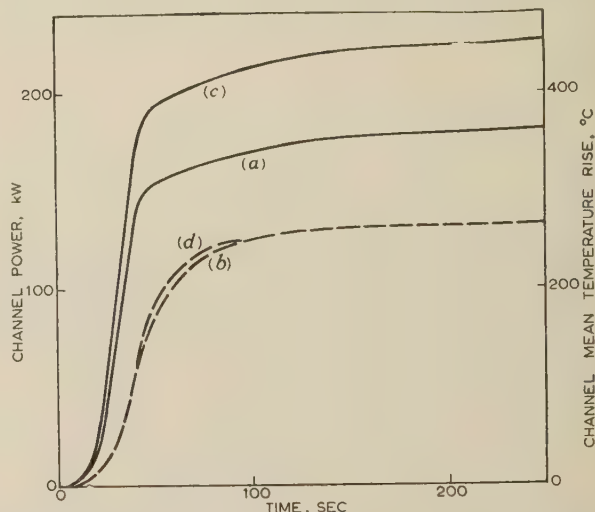
Fig. 18.—Transient power and temperature variations with time.

    (a) Power in channel r^*.
    (b) Mean temperature in channel r^*.
    (c) Power in central channel.
    (d) Mean temperature in central channel.

The corresponding channel mean temperature transient in the central channel is plotted from eqn. (38), where the value of τ is that corresponding to central-channel conditions, and the equation becomes

$$d\bar{T}/dt = 0 \cdot 129P - 0 \cdot 108\bar{T}_u \quad . \quad . \quad . \quad (42)$$

The transient of maximum temperature occurring in the central channel, from eqn. (20), is $1 \cdot 22$ times that plotted as curve (d) in Fig. 18.

The validity of using the ratio obtained from the steady-state distribution may be checked by evaluating the distribution parameter $\gamma z/\nu$. For a diameter of fuel rod of $1 \cdot 1$ in and flow area in channel of 10 in² (typical dimensions), the value of γ, assuming CO_2 at 100 lb/in², is $67 \cdot 2$. ν = Coolant velocity × uranium thermal time-constant = $57 \cdot 2 \times 11 \cdot 4 = 650$ ft. From Fig. 8, at the point of maximum uranium temperature, $z = 15$ ft, approximately. Therefore, $\gamma z/\nu = 1 \cdot 55$.

From Fig. 16, the transient deviation from the steady-state distribution is judged to be insignificant.

THE DESIGN, PERFORMANCE AND USE OF FISSION COUNTERS

By W. ABSON, B.Sc., Associate Member, P. G. SALMON and S. PYRAH.

SUMMARY

The basic design criteria for electron-collection fission counters are discussed; data are presented concerning the effect of thickness of fissile material on sensitivity and the effect of electrode spacing and gas pressure on pulse height. High detection efficiencies are obtained with electrode coatings of 1 mg/cm^2, and thicknesses up to 2 mg/cm^2 may be used when high sensitivities are required. When argon is used as the filling gas the product of gas pressure and electrode spacing (pd, in atmosphere-cm) should lie in the region 0·4–1·0, the higher value being preferable if it does not give rise to difficulties in ionization collection. The operating level of the discriminator bias is usually determined by the effect of α-activity in the fissile material. Typical values of the operating bias level at which a high detection efficiency is possible correspond to 5–20 MeV particle energy.

The designs of several types of fission counter are described, illustrating experimental applications and the use of this type of detector in reactor instrumentation systems.

(1) INTRODUCTION

The detection of high-energy ionizing particles, resulting from the neutron-fission process in a thin layer of fissile material on the plates of an ionization chamber, may be used in the measurement of fission cross-sections and neutron-flux levels. In the former type of measurement, very thin layers of fissile material are used (0·1 mg/cm^2 or less), so that all fission events give rise to an ionizing track (from one of the two fission fragments) with a total energy dissipated greater than that due to α-particle emission. For relative flux measurements, greater thicknesses of coating may be used (1–2 mg/cm^2), in order to obtain a greater sensitivity to neutrons with some sacrifice of the efficiency with which fissions are detected. For slow-neutron detection and maximum sensitivity, coatings of uranium 235 are used, whilst uranium 238 or thorium 232, which have neutron-fission thresholds at about 1 MeV, may be used for fast-neutron measurements.

Many types and sizes of fission counters, for use in nuclear physics measurements and for reactor control applications, have been described in the literature.[1,2] The counters to be described were developed for measurements in reactor core assemblies and for use in reactor instrumentation systems, and emphasis has been placed on ruggedness of design. The various types described also serve to illustrate the range of application of fission counters, from large multiple-plate high-sensitivity counters for low flux level measurements, to small counters required for flux scanning with a high spatial resolution. An important characteristic of fission counters is their low sensitivity to γ-radiation. For example, it is possible to operate a fission counter with a cathode area of the order of 100 cm^2 at γ-radiation levels of 10^5 rads/h and still obtain discrimination between neutron and γ-radiation. This is to be compared with an upper limit of the order of 100 rads/h for a counter of similar size using the n, α reaction in boron 10 (e.g. a boron-trifluoride proportional counter). Fission counters may therefore be the only type of detector which can be used in the measurement of low power

levels in small-size reactors, where it is necessary to place the detector near the reactor core yet not possible to provide additional shielding from fission-product γ-radiation. These detectors may also be required to operate at high temperatures. Most of the counters described are suitable for operation up to 200° C; some brief details are also given of counters being developed for operation at 500° C.

In Section 2 the basic design criteria for fission counters are listed and discussed. This Section also includes design data obtained in the course of recent counter developments. Some specific designs and their application are described in Section 3.

(2) GENERAL DESIGN CRITERIA

Fission counters are generally used as electron-collection ionization chambers in order to obtain a short response time. The fissile material is applied to the negative electrode only, the gas pressure being adjusted so that the centres of the majority of the ionized tracks are close to the negative electrode; the voltage induced on the anode due to the collection of the primary electrons thus approaches the maximum possible value. The gas pressure, p, and the electrode spacing, d, will also govern the ionization-collection efficiency and the response-time of the counter. Data on the dependence of sensitivity on fissile coating thickness, and the effect of gas filling on pulse height, ionization-collection efficiency and on transit times are given below.

(2.1) Sensitivity: Dependence on Thickness of Fissile Material Coating

The two main groups of fission particles (the heavy and light fragments) have most probable energies of about 65 and 95 MeV, and corresponding ranges in air at s.t.p. of 1·9 and 2·5 cm. These ranges are approximately the same as for α-particles of 3–4 MeV energy. In the oxide coatings (e.g. U_3O_8) used in fission counters, the mean estimated[1] range of fission fragments is 10 mg/cm^2. An approximate estimate of the fraction, F_e, of fission fragments emerging from a flat coating of thickness t is given by

$$F_e = 1 - \frac{t}{2(R_0 - R_E)} \quad . \quad . \quad . \quad . \quad (1)$$

where F_e is the fraction of particles emerging with a remaining range equal to or greater than R_E, and R_0 is the maximum range.[1] In order to detect a fission event and distinguish it from an α-particle emission, the energy of the fission fragment emerging from the coating must be greater than the maximum energy of the α-particles (4–5 MeV). Where the total α-activity in the coating is high compared with the fission rate, it is necessary to set the discriminator bias at a level corresponding to a much higher particle energy, in order to distinguish between fission particle pulses and several α-pulses occurring within the response-time of the counter/amplifier system. This effect of build-up of α-pulses is discussed in more detail in Section 2.5. Assuming for the present that the remaining energy must be 10 MeV, the corresponding value for R_E is about 4 mg/cm^2. Substituting $R_0 = 10$ mg/cm^2 and $R_E = 4$ mg/cm^2 in eqn. (1) gives $F_e = 0·99$

The authors are at the U.K.A.E.A. Atomic Energy Research Establishment, Harwell.

for $t = 0.12$ mg/cm^2. Coatings of about 0.1 mg/cm^2, and preferably less, are therefore necessary for absolute measurements of neutron flux and for fission cross-section measurements. For relative flux measurements, greater thicknesses of coating are permissible; e.g. for 1 mg/cm^2 the detection efficiency given by eqn. (1) is approximately 90%.*

Experimental data for U_3O_8 coatings of 0.12, 1.0, 2.1 and 3.2 mg/cm^2 deposited over an area of 5 cm^2 of the cathode of a parallel-plate counter are given in Figs. 1 and 2. The counting-
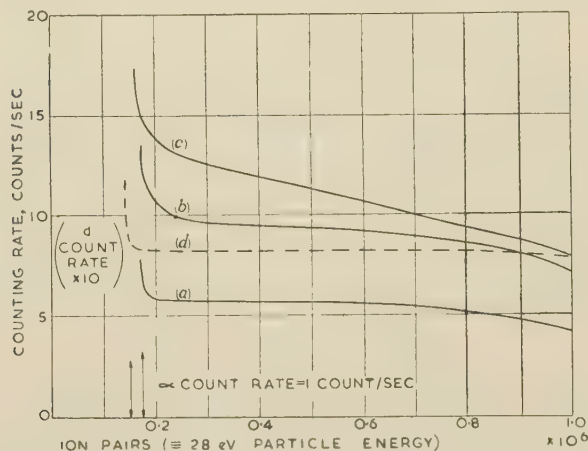
Fig. 1.—Fission-counter sensitivity: dependence on thickness of fissile material coating.

Discriminator-bias curve : for
(a) 1.0 mg/cm^2 cathode coating.
(b) 2.1 mg/cm^2 cathode coating.
(c) 3.2 mg/cm^2 cathode coating.
(d) 0.12 mg/cm^2 cathode coating.
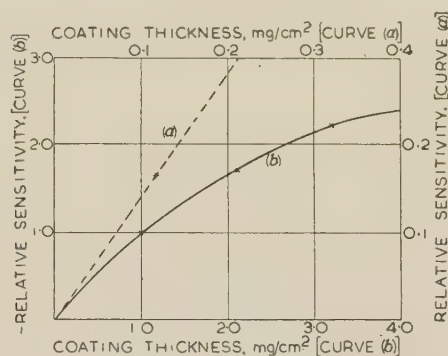
Fig. 2.—Sensitivity as a function of U_3O_8 coating thickness at a discriminator-bias level of 0.36×10^6 ion pairs or 10 MeV.

rate/discriminator-bias measurements shown in Fig. 1 were made with four counters placed in turn at a fixed distance from a radium-beryllium source surrounded by a paraffin-wax moderator. The electrode spacing was in each case 0.3 cm and the gas filling was argon and 2% nitrogen at a pressure of 5 atm. With this gas filling the shape of the discriminator bias characteristic (i.e. the pulse-height distribution) is substantially independent of gas pressure and is mainly determined by absorption effects within the fissile material coating.

The observed increase in sensitivity with increasing thickness of coating is somewhat less than that predicted by eqn. (1), which does not take account of the spread in energy and range of the fission fragments. Assuming that the detection efficiency for the very thin coating (0.12 mg/cm^2) is about 99%, the efficiencies at 1.0 and 2.0 mg/cm^2 are 71 and 59%. The efficiency

* The detection efficiency is defined here as the ratio of the observed counts to the total number of fission events.

for other thicknesses of coating can be obtained from Fig. 2, which shows the variation of sensitivity with thickness at a discriminator bias level equivalent to 10 MeV particle energy. At higher operating bias levels, which may be necessary to avoid α build-up effects, the graph of sensitivity as a function of coating thickness will be more sharply curved. For thicknesses greater than about 2 mg/cm^2 there is a marked increase in the slope of the discriminator-bias curve and the increase in sensitivity is far from commensurate with the increase in the amount of fissile material. There is thus little practical advantage to be gained in the use of coatings of thickness greater than 2 mg/cm^2. Similar results have been quoted by Baer and Swift.[3]

The effective sensitivity of a fission counter may also be reduced by absorption of neutrons in the materials of the counter. For example, for multiple-plate counters of the type shown in Fig. 7, which are made from steel tubes approximately 0.05 in thick, the reduction in thermal-neutron sensitivity due to absorption effects is 10–15%.

It is convenient to refer to the pulse amplitude or the discriminator-bias level in terms of an equivalent particle energy W_e, or an equivalent collected charge Q_e, where Q_e may be expressed in coulombs or ion pairs. For an energy of W electron-volts $\times 10^6$ dissipated in an ionized track in argon, the number of ion pairs formed is $(W \times 10^6)/28$, and the total charge Q of the electrons produced is $W \times 5.7 \times 10^{-15}$ coulombs. For a parallel-plate system, the voltage pulse induced at the anode due to collection of these electrons is $(Q/C)(l/d)$, where C is the total capacitance at the anode, l is the mean distance travelled by the electrons, and d is the electrode spacing. Thus $Q_e = Q(l/d)$ or $W_e = W(l/d)$. If the collection time is short compared with the amplifier response-time the output pulse amplitude is then $(Q_e/C)A$, where A is the amplifier gain for a step-function pulse. In Fig. 1, and other discriminator-bias curves in the paper, the discriminator scale is given in units of 10^6 ion pairs, which is equivalent to a charge Q_e of 1.6×10^{-13} coulombs or a particle energy (in argon) of 28 MeV. The input charge sensitivity of a counter and amplifier system may be measured by injecting a voltage pulse, V, through a small series capacitance, C_s, at the counter anode. If C_s is very small compared with the input capacitance C, the input charge is C_sV and the pulse-generator output may be calibrated in terms of ion pairs or coulombs or particle energy.*

The choice of the operating level for the discriminator bias is determined by two considerations: the need to obtain a high detection efficiency, and the need to reject pulses due to α-activity in the fissile material. When the total α-activity is comparable with the observed fission rate, it is permissible to use bias settings only a little higher than the maximum energy of the α-emission (4–5 MeV), but for higher activities the probability of two or more α-pulses occurring within the response-time of the system will increase, and higher bias settings will be necessary in order to reject these multiple pulses. In Fig. 1 curve (c), for example, which refers to a counter with an α-emission of about 10^4 per second, it is necessary to use a bias level equivalent to about 0.19×10^6 ion pairs, or about 5.3 MeV particle energy, to reduce the α-response to 1 count/sec. Some further examples of the effect of multiple α-pulses are considered in Section 2.4 on counter and amplifier response times.

(2.2) Effect of Electrode Spacing and Gas Filling on Pulse Amplitude

Fig. 3 shows a series of curves for counting rate as a function of discriminator bias for a parallel-plate counter with an electrode

* The A.E.R.E. pulse generator type 1405 may be used over the range 0–10^6 ion pairs. A series capacitor (C_s) of 0.16 pF is used, i.e. a charge of 10^6 ion pairs corresponds to a voltage pulse, V, of 1 volt.
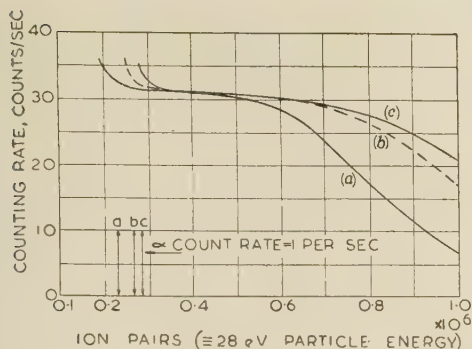
Fig. 3.—Effect of gas filling pressure on discriminator-bias curves.

Electrode spacing = d = 0·1 cm.   Cathode coating: 1 mg/cm^2.
$T_1 = T_2 = 0·08$ microsec.

Total α-emission $\simeq 10^5$ per second.
   (a) Argon pressure = 4·5 atm.
   (b) Argon pressure = 7·0 atm.
   (c) Argon pressure = 10·0 atm.

spacing of 0·1 cm filled with argon and 2% nitrogen at various pressures. The thickness of the coating (U_3O_8) on the negative electrode was in this case 1 mg/cm^2. As the gas pressure is increased, the observed pulse amplitudes increase, owing to a rise in the energy expended in the ionized tracks and an increase in the mean distance through which the electrons move before collection at the anode. There is no critical optimum value of pd. The amplitude of the larger pulses continues to rise up to $pd = 1·0$ atm-cm, but the detection efficiency and the slope of the discriminator-bias curve at typical operating levels are approaching constant values at lower pressures. The improvement in the fission count characteristics between curves (b) and (c) (for $pd = 0·7$ and 1·0) is offset by an increased α-response. It may sometimes be preferable to use a still lower value of pd, e.g. 0·45, as in curve (a), with some increase in the slope of the bias characteristic, in order to reduce the polarizing voltage required to achieve an adequate ionization-collection efficiency.

(2.3) Ionization-Collection Characteristics

In an electron-pulse counter the pulse amplitude is reduced if electrons recombine with positive ions or if they become attached to neutral molecules to form heavy slow-moving ions. Attachment of electrons to oxygen impurities is the most important practical effect, and Facchini and Malvicini[4] have shown that the addition of 1–2% nitrogen to argon reduces the magnitude of this effect at values of E/p (ratio between electric force and gas pressure) likely to be used in electron-pulse counters. The attachment coefficient for oxygen is a function of the electron energy, and with 'pure' argon, containing traces of oxygen, the pulse height increases at first with increasing electric force and then decreases owing to an increase in the attachment coefficient. This resonance effect is avoided in an argon-nitrogen mixture, as a much higher value of E/p is required to attain electron energies in the resonance region. A further advantage of argon-nitrogen mixtures is that the electron drift velocity is considerably increased, thus leading to shorter transit times in the counter.

In a pulse counter which exhibits a plateau region on its discriminator-bias characteristic, it is not essential to achieve 100% ionization-collection efficiency, as a change in pulse amplitude due to variation in applied voltage will not produce a corresponding change in counting-rate in the plateau region. A satisfactory operating voltage can be determined by operating the counter with the discriminator bias level set at a point below the knee of the bias curve and adjusting the inter-electrode voltage until the counting rate is within 10% of the maximum possible at

this bias setting. At this voltage, V_p, the counting rate in the plateau region will usually be within 1% of the maximum value. Fig. 4 shows the variation of counting-rate with polarizing voltage (near 50% maximum counting-rate) for various combinations of electrode spacing and gas pressure. It will be seen that as the spacing is increased it is possible to use greater values of pd for a given value of V_p.
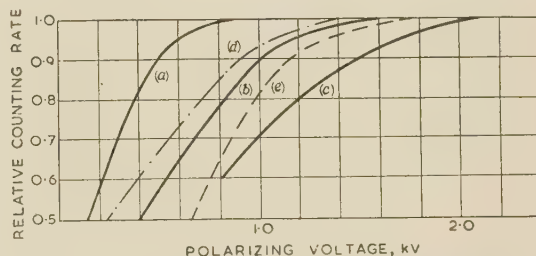
Fig. 4.—Fission-counter saturation characteristics.

Counting-rate as a function of polarizing voltage, below knee of bias curve.
$T_1 = T_2 = 0·08$ microsec.

   (a) Pressure p = 4·5 atm (argon).   Electrode spacing d = 0·1 cm.
   (b) Pressure p = 7·0 atm (argon).   Electrode spacing d = 0·1 cm.
   (c) Pressure p = 10·0 atm (argon). Electrode spacing d = 0·1 cm.
   (d) Pressure p = 4·5 atm (argon).   Electrode spacing d = 0·2 cm.
   (e) Pressure p = 2·0 atm (argon).   Electrode spacing d = 0·7 cm.

The results quoted in Fig. 4 were obtained with low fission counting-rates, of the order of 1 count/sec per square centimetre of cathode area. The collection characteristics are, however, not markedly affected at increased counting-rates. Small-area counters of the type shown in Fig. 9 (coated cathode area = 3·6 cm^2) have been used at fission counting-rates of the order of 10^5 counts/sec per square centimetre of cathode. V_p is increased by some 10–20%, and the approach to 100% collection is slower than in the case of low counting-rates. This is presumably due to the effects of positive-ion space charge producing a reduction in field strength near the anode. For most applications of pulse-type fission counters, the fission rate per square centimetre of cathode will be less than 10^5 per second, and the collection characteristics can be considered to be independent of counting-rate.

(2.4) Transit Times: Counter and Amplifier Response Times

The maximum rise-time of the voltage pulse in a parallel-plate fission chamber is given by d/v, where v is the electron drift velocity. For electrons in argon and 1% nitrogen mixture, v increases with E/p in an approximately linear fashion to about 2·2 cm/microsec for $E/p = 700$ volts/cm per atmosphere, and then tends towards a constant value.[5] It can be seen from the results quoted in Fig. 4 that the values of E/p likely to be used for fission counters with electrode spacings between 0·1 and 0·7 cm lie in the range 600–1 400; over this range v does not change appreciably and has an average value of about 2·0 cm/microsec.* Thus the estimated pulse rise-times vary from about 0·05 microsec for 0·1 cm spacing, to about 0·35 microsec for 0·7 cm spacing.

The relative voltage amplitude of fission pulses for counters with 0·1 and 0·2 cm spacings, when used with an A.E.R.E. type 1049C amplifier,† with various settings of the differentiating and integrating time-constants (T_1 and T_2), are shown in the full-line curves of Fig. 5. The relative amplitudes shown here were obtained from the discriminator bias settings required to give half-maximum counting-rate.

---

* This is some five times the value of v in pure argon for the same value of E/p.
† The type 1049C pulse amplifier consists of a head amplifier unit with a gain of about 100 and a main amplifier with a maximum gain of about 10^4. The main unit is provided with attenuator controls covering a range 1–100 (40 dB) and with controls for adjusting RC-type differentiating and integrating time-constants.
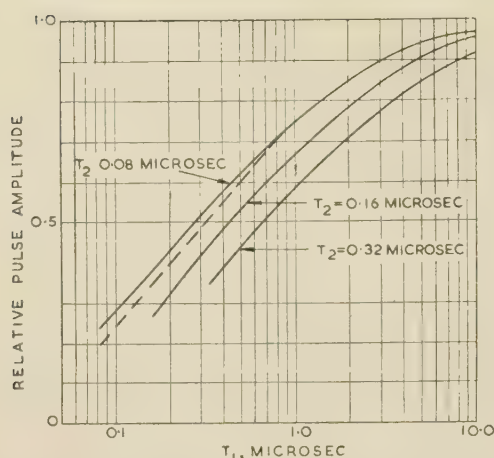
Fig. 5.—Effect of differentiating and integrating time-constants, T_1 and T_2, on amplitude of fission-counter pulses.

——— Counters with 0·1 and 0·2 cm electrode spacing.
- - - - Counter with 0·7 cm electrode spacing.

The effect of amplifier time-constants on pulse amplitude and pulse shape, when the input signal is a step function with a uniform rate of rise (as in parallel-plate electron-collection pulse counters), has been considered in detail by Gillespie.[6] It can be shown that with input rise-times which are smaller than or comparable with the amplifier integrating time-constant, the output pulse amplitude and shape are practically the same as for an infinitely short rise-time. When a pulse-generator signal with a rise-time of 0·1 microsec was injected into the input circuit of the type 1049C amplifier used for the fission counter experiments, curves identical to those of Fig. 5 were obtained, showing that the rise-times for 0·1 and 0·2 cm spacings are of the order of 0·1 microsec. The curve obtained for $T_2 = 0·08$ microsec with a 0·7 cm spacing fission counter (the dotted curve of Fig. 5) showed a slightly greater reduction in pulse height with decreasing values of T_1, indicating a higher value for the rise-time of this counter.

The theoretical estimate for the reduction in amplitude at $T_1 = T_2$ (for an input step-function of rise-time $T < T_2$) is $1/\varepsilon = 0·37$. Fig. 5 shows reductions to 0·22, 0·26 and 0·34 for values of $T_1 = T_2$ of 0·08 microsec, 0·16 microsec and 0·32 microsec. Two factors contribute to this effect. At low settings of the T_1 control, there is some attenuation of signal because of stray capacitance across the resistor in the RC circuit. At the lower settings of the integrating time-constant control the effective value of T_2 is also governed by other circuit parameters controlling the upper limit of the amplifier frequency response.

The resolving time T_r of the system, i.e. the period after one pulse has triggered the discriminator before the signal amplitude has fallen below the triggering level and a second pulse may be registered, is governed by the rise-time of the counter, T, and by T_1 and T_2. Gillespie has shown that, for the case of exponential time-constants in the amplifier, T_r is given by

$$T_r = \frac{V_0}{V_m} T_1 \quad . \quad . \quad . \quad . \quad . \quad (2)$$

where V_m is the measured voltage amplitude and V_0 is the maximum observable amplitude, i.e. with $T_1 \gg T_2$.

Although this expression will not apply strictly to the practical case of a counter producing pulses with a wide distribution in amplitude or for effects in the amplifier due to stray capacitance, etc., it is still a useful guide in determining the optimum settings of T_1 and T_2. Thus, when a short resolving time is required, T_2

should be set to a value comparable with the pulse rise-time in the counter, as this ensures the minimum value of V_0/V_m for any subsequent setting of T_1. T_1 should then be set to give the best compromise between the values of T_r and V_m. When T_1 is comparable with T_2, T_r does not change rapidly as T_1 is varied, and it is often preferable to use values of T_1 equal to 2 or 3 times T_2 and obtain an increased output pulse at the expense of a slight increase in resolving time. Using eqn. (2) and the results of Fig. 5 (curve for $T_2 = 0·08$ microsec), $T_r = 0·33$ microsec at $T_1 = 0·08$ microsec and is increased to 0·43 microsec at $T_1 = 0·16$ microsec. The increase in resolving time is some 30%, but the corresponding increase in pulse amplitude is 60%. This point is illustrated further in the next Section in relation to the α build-up effect.

(2.5) High α-Activity and High Fission Counting-Rates

One of the limitations to the amount of fissile material which may be used in a counter is the total α-particle emission from the coating. The probability P_n that n α-emissions may occur within the resolving time T_r is given by

$$P_n = \frac{(\lambda T_r)^n}{n!} \varepsilon^{-\lambda T_r} . \quad . \quad . \quad . \quad . \quad (3)$$

where λ is the rate of emission of α-particles. Multiple pulses of n unresolved events will then occur at the rate of P_n/T_r per second. Thus, for $T_r = \frac{1}{3}$ microsec and $\lambda = 3 \times 10^6$ per second, $\lambda T_r = 1$ and $P_n = 1/(n!\varepsilon)$. For $n = 10$, $P_n/T_r \simeq 3$, i.e. 10 events will occur within the resolving time at a rate of 3 per second. The equivalent bias level at which 3 counts/sec will be observed will then be 10 times the average amplitude of the individual α-pulses.

An example of a large α build-up effect is found in the large-area counter shown in Fig. 7. The total α-emission, with 1 500 cm² of cathode area coated with U_3O_8 highly enriched in uranium 235 at a thickness of 2 mg/cm², is of the order of 3×10^6 per second. Fig. 8 shows the counting-rate due to α-activity as a function of discriminator bias level, and it will be seen that a discriminator level equivalent to $4·5 \times 10^5$ ion pairs ($\equiv 13$ MeV particle energy) is required in order to reduce the counting-rate to 3 per second. On the basis of the simple theory outlined above, which for this case gives 10 α-emissions per resolving time, the average α-energy per pulse is thus 1·3 MeV. This is a reasonable figure for a 2 mg/cm² coating.

The effect on the α build-up in this counter of different combinations of T_1 and T_2, in a type 1049C amplifier, is shown in Table 1. Values for the discriminator level, at which the α-response is reduced to 1 count/sec, are given in ion pairs.

Table 1

MAGNITUDE OF α BUILD-UP (EQUIVALENT ION PAIRS) FOR A COUNTING RATE OF 1 PER SECOND AND A TOTAL α EMISSION OF 3×10^6 PER SECOND

Integrating time-constant, T_2	Differentiating time-constant, T_1		
	microsec 0·08	microsec 0·16	microsec 0·32
microsec 0·08	$4·9 \times 10^5$ (1·0)	$5·3 \times 10^5$ (1·7)	$5·8 \times 10^5$ (2·5)
0·16		$6·3 \times 10^5$ (1·2)	
0·32			$7·2 \times 10^5$ (1·5)

Although the α build-up effect has its least value at the shortest time-constant settings ($T_1 = T_2 = 0.08$ microsec), this is only achieved at the expense of a reduction in output signal voltage. The relative signal amplitudes are shown in parentheses in Table 1, with an arbitrary value of 1.0 for $T_1 = T_2 = 0.08$ microsec. The signal amplitude may be increased by a factor of 2.5 (with $T_1 = 0.32$ microsec and $T_2 = 0.08$ microsec) for an increase in α build-up of only 20%. There is clearly no virtue in using higher values of integrating time-constant.

The percentage counting losses due to coincidence of fission pulses during the resolving time is given approximately by $\bar{n}T$, where $\bar{n}$ is the mean counting-rate. With the amplifier time-constants set at $T_1 = T_2 = 0.08$ microsec, and a counter pulse rise-time of the same order of magnitude, $T_r \simeq \frac{1}{3}$ microsec, a 10% counting losses should therefore occur for $\bar{n} = 3 \times 10^5$ per second. This is in agreement with experimental results obtained using two counters with sensitivities differing by a factor of 100 and moving them together into a region of increasing neutron flux. The more sensitive counter exhibits counting losses of about 10% at 3×10^5 counts/sec.

baking operations. The amount deposited on an electrode may be assessed by weighing, but it is usually necessary to pre-bake the electrode until it attains constant weight. Alternatively, a given amount of material may be dispensed and used up completely on an electrode assembly. In this case it is necessary to ensure that no appreciable amount of material is retained in brushes or on the walls of containers. When stainless steel is used as the electrode material the baking process may be carried out in air, as the weight of oxide formed is negligible (compared with 1 mg/cm^2) and it is in any case a hard adherent film. When aluminium is used as the electrode, it is preferable to use an inert gas in the baking oven to avoid formation of oxide films. Similar methods have been used for preparing other electrode coatings such as plutonium 239, uranium 233 and 238 and thorium 232.

(3.1) Medium-Size Multiple-Plate Fission Counter

Fig. 6 shows the cross-section of a multiple-plate chamber with coaxial cylindrical electrodes with a 0.1 cm spacing, and a total cathode area of 165 cm^2 (A.E.R.E. type B165 counter). The length of the coated area is 2 in and the electrode assembly is
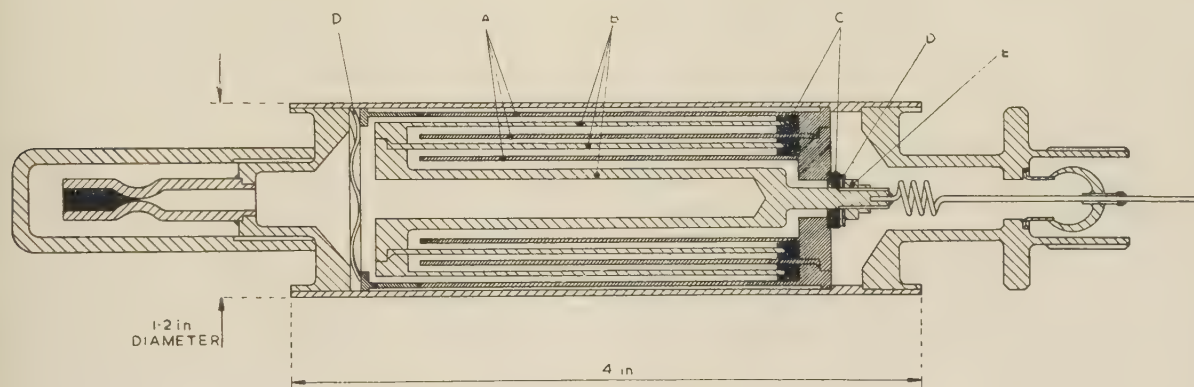
Fig. 6.—Multiple-plate fission counter type B165.

Coatable cathode area = 165 cm^2.   Electrode spacing = 0.1 cm.

A Coated cathode tubes.
B Anode.
C Quartz insulators.
D Nimonic springs.
E Fixing nut for anode/cathode assembly.

(3) EXAMPLES OF FISSION-COUNTER DESIGNS AND THEIR APPLICATIONS

Uranium coatings with various degrees of enrichment in uranium 235 are used for slow neutron-flux measurement, as this type of material has a high fission cross-section coupled with low specific α-activity. The electrodes may be coated by means of successive applications of a solution of uranyl nitrate in alcohol and acetone. The method has been generally described in the literature[1] and will only be outlined here. A typical solution concentration for coatings of 1 mg/cm^2 is 5 mg/ml of uranium with 2–3$\%$ by weight of nitro-cellulose lacquer (e.g. Zapon). The solution is painted on to the electrode using a brush or felt pad and is allowed to dry at room temperature. The electrode is then baked for about 15 min at 500° C; the Zapon binding agent is burnt off and the uranyl nitrate is reduced to the oxide U_3O_8, which remains as a tough coherent film. The maximum amount of material which may be applied per coating in this way is about 0.1 mg/cm^2. Attempts to increase this by the use of more concentrated solutions or applying more solution per coating usually lead to the formation of non-adherent layers. The total required thickness (up to 1 or 2 mg/cm^2) is obtained by successive painting and

contained within a cylindrical envelope of 1.2 in diameter. The electrodes and outer tube are all made from stainless steel of wall thickness 0.02 in. Five of the cathode surfaces may be coated (the inside and outside of two cylinders and the inside surface of the outer one) giving a total area of 165 cm^2. After the coating process the cathode and anode coaxial cylinders are fitted together as a complete assembly, the various parts being located with respect to one another by means of grooves in the surface of the quartz insulators. A Nimonic-alloy spring D is used under the fixing nut E to take up any relative expansions which may occur during outgassing processes or operation of the counter at high temperatures. The assembly of the outer envelope is completed except for the welding of the bottom end-cap; the electrode assembly is then fitted into the envelope, the anode-lead wire being threaded through the central tube of the glass/metal (or ceramic/metal) seal. The remaining operations before pumping and filling are then limited to the welding of the end-cap and the brazing of the anode wire to the central tube of the glass/metal seal. It is desirable to arrange the design of fission counters so that final assembly operations are as simple as possible, since in cases where the coating has a high specific α-activity it is necessary to do the whole of this work under dry-box conditions. The
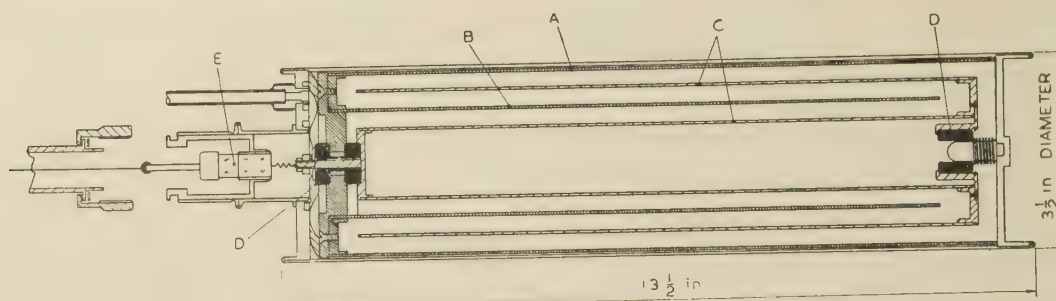
Fig. 7.—High-sensitivity fission counter.

Coatable cathode area = 1 500 cm². Electrode spacing = 0·7 cm.

A Coated cathode tube.
B Coated cathode tube.
C Anode.
D Quartz insulators.
E Ceramic/metal seal.

counter may be fitted with a concentric air-spaced lead, to provide a low-capacitance connection to a pulse amplifier.

The type B165 counter and a smaller counter of similar design (type A35, cathode area 35 cm²) have been extensively used, with a variety of fissile material coatings, for neutron-flux measurements in experimental reactor core assemblies. The B165 is used with a 1 mg/cm² coating of U_3O_8, highly enriched in uranium 235, in the low-power-level control instrumentation of the Lido light-water reactor at the A.E.R.E. The sensitivity for thermal neutrons for a fission counter with this coating is about 0·12 counts/sec per unit neutron flux, and it is suitable for flux measurements up to about 10^6 neutrons/cm²/sec. When used with short amplifier time-constants ($T_1 = T_2 = 0·08$ microsec) the response to γ-radiation levels of 3×10^4 rads/h is still small compared with the α-response level, and this size of counter could still be used to discriminate between neutron flux and γ-radiation at considerably higher γ dose-rates.

(3.2) Large-Area High-Sensitivity Fission Counter

The fission counter shown in Fig. 7 has been designed for use in a reactor control-instrumentation system for neutron-flux level indication and for operation of a reactor-period meter over the range 25–10⁵ neutrons/cm²/sec. A 2 mg/cm² coating of U_3O_8 (highly enriched in uranium 235) is used over an area of 1 500 cm² to give a thermal-neutron sensitivity of about 1·7 counts/sec per unit flux. The coating extends over three of the cylindrical-cathode surfaces, the inner of tube A and both surfaces of tube B. The method of assembly is similar to that described for the type B165 counter, the cathode and anode electrodes being fixed together before assembly inside the main envelope. The counter is constructed from stainless steel, the wall thickness of the electrode tubes and the outer case being 0·05 in.

The discriminator-bias curve for neutron fission and for the α-particle emission from the coating are shown in Fig. 8. The operating bias level must be greater than about $5·5 \times 10^5$ ion pairs (15 MeV) in order to operate at low neutron counting rates and obtain discrimination from the α-particle response. This is approaching a practical upper limit of operating bias for a counter with a 2 mg/cm² coating. The slope of the discriminator-bias curve at this point is about 2·5% for a 10% change in bias level and is increasing rapidly with increasing bias.

The application for which the counter was designed provides an example of the limitations of total capacitance and the compromises in design which are necessary in order to achieve a given sensitivity. A typical figure for valve noise in a pulse amplifier with time-constants T_1 and T_2 set at 0·08 microsec is 20 μV. (This refers to the equivalent input discriminator-bias
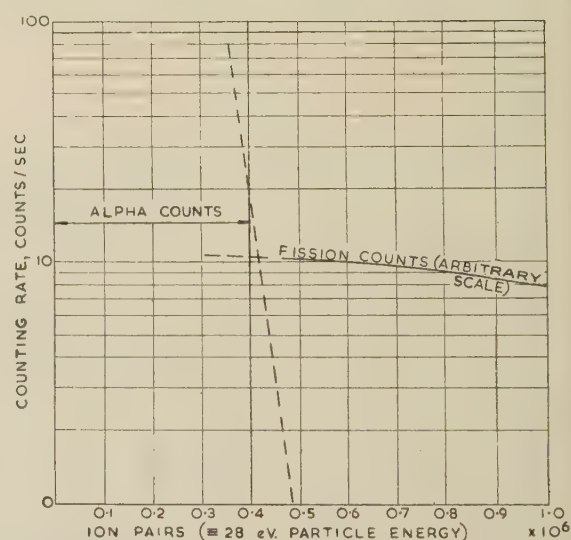
Fig. 8.—Fission and α-bias curves for high-sensitivity counter of Fig. 7.

Gas filling, argon + 2% nitrogen at 2 atm.
Total α emission ≃ 3×10^6 per second.
$T_1 = T_2 = 0·08$ microsec.

level at which a counting rate of a few counts per second will k observed.) In a control-instrumentation system it is desirab to set the input discriminator level at some 10 times this valu in order to minimize other sources of spurious pulses (e.g. due electrical interference); in the case under discussion a furthe reason for using a high input level of discrimination is that th peak amplitude of valve noise pulses should be small compare with the magnitude of the α build-up. Thus, if an equivale particle energy of 15 MeV ($\equiv 5·5 \times 10^5$ ion pairs or $8·2 \times 10^{-}$ coulombs) is to produce a voltage pulse greater than 200 μV the maximum permissible value for the input capacitance 400 pF. In the reactor application, the capacitance of the con necting leads to the pulse amplifier is about 200 pF, leaving maximum permissible counter capacitance of 200 pF. The sens tivity requirement (1·5–2 counts/sec per unit neutron flux) ma it necessary to use some 3 g of uranium 235. With a coatin thickness of 1 mg/cm² the area of coating would then be 3 000 cm and an electrode spacing of about 1·2 cm would be required t limit the capacitance to 200 pF. Because of overall-size limita tions a coating thickness of 2 mg/cm² was chosen, with a coate area of 1 500 cm² and an electrode spacing of 0·7 cm.

(3.3) **Small Fission Counter for Flux-Scanning Applications**

An example of a small fission counter for the measurement of
flux distribution with a high spatial resolution is shown in Fig. 9.
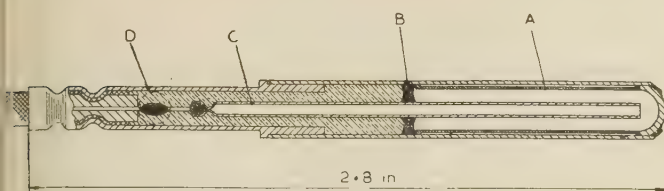
Fig. 9.—Small-size flux-scanning counter.
Diameter 1/4 in approx.

A Coated cathode tube.
B Glass/metal seal.
C Anode.
D Extension tube connecting to flexible cable.

This counter was developed to enable flux measurements to be
made along a $\frac{1}{4}$ in diameter pipe attached to fuel elements of
the type used in the Dido heavy-water reactor.   The outside
diameter and the overall length of the counter envelope were
chosen so as to allow the counter and its cable connector to be
passed along bends in the pipe of 25 in radius.   The outer case
of the counter and the anode are made from Nilo-K* tubing
(0·01 in thick) so that an internal glass/metal seal, B, can be
made between these two components.   The fissile material
coating is applied to a separate cathode liner, A.   After fitting
this liner and the rounded end-cap, the counter is evacuated
and then filled with the argon-nitrogen mixture, through the
anode tube C, which is then crimped and sealed.   The polythene-
insulated connector lead is secured to the counter body by
forming a rolled groove in the tube D, which compresses the
outer braided conductor of the cable into the polythene core.
The soldered joint between the centre conductor of the cable
and the counter anode lead is made using a small soldering bit
through a hole in D.   The inside of D is then filled with poly-
thene by injection through this hole.

The cathode liner may be coated over a length of up to 2·5 cm,
giving a maximum coated area of 3·6 cm². When coated with
U_3O_8 highly enriched in uranium 235 (93%) at a thickness of
1 mg/cm², the sensitivity to thermal neutrons is approximately
3 × 10⁻³ counts/sec per unit flux.   For an upper limit of
counting-rate of 3 × 10⁵ counts/sec, the corresponding upper
limit of flux is then 10⁸ neutrons/cm²/sec.   A number of these
counters have been used with this sensitivity, and also with a

reduced weight of fissile material, to reduce the sensitivity by a
factor of about 100, so that they may be used for flux measure-
ments up to 10¹⁰ neutrons/cm²/sec.

Fig. 10 shows typical counting-rate/discriminator-bias curves
for this type of counter for U_3O_8 coatings of 1 mg/cm² and
approximately 10⁻² mg/cm².   The improved bias characteristics
obtained with the thin coating are due to the reduced absorption
of fission-particle energy within the coating.   The pulse-height
distribution in this counter is affected by the curvature of the

* An alloy of nickel, cobalt and iron.

Fig. 10.—Discriminator-bias curves for $\frac{1}{4}$ in diameter fission counter.
Gas filling, argon + 2% nitrogen at 4·5 atm.
(a) 1 mg/cm² U_3O_8.
(b) Approximately 0·01 mg/cm² U_3O_8.

cathode, when a gas filling pressure of 4·5 atm is used.   Fission
particles with a small angle of emergence from the cylindrical-
cathode surface will strike the cathode before dissipating the
whole of their energy in the gas.   The effect is more marked
for the 1 mg/cm² coating, as a greater proportion of particles
emerging at a small angle to the cathode will have dissipated an
appreciable energy in the coating and will have a significantly
lower specific ionization.   This accounts for the larger slope
on the bias curve for a 1 mg/cm² coating in the $\frac{1}{4}$ in counter
compared with the larger-diameter or flat-plate type of counter.
The slope can be reduced by using a higher pressure for the gas
filling, but this leads to an inconveniently high polarizing voltage.
With the small clearances in the cable connections to this counter,
the lowest possible voltage operation is required.   With 4·5 atm
filling it can be operated satisfactorily at about 500 volts, and is
free from spurious pulses.

(3.4) **High-Temperature Operation of Fission Counters**

Fission counters of the type shown in Fig. 6 and using
glass/metal seals have been operated satisfactorily at tempera-
tures up to 200° C.   Above this temperature ionic conduction

Fig. 11.—Coaxial-tube fission counter with ceramic/metal insulator support and seal.
A Coated cathode tube.   B Anode.   C Ceramic/metal seal.   D Ceramic spacer.

processes begin to occur in the glass, and this gives rise to
spurious pulses.

Ceramic/metal seals using Nilo-K tubing sealed to alumina
have been tested to 500° C and found to be free from spurious
pulses with applied voltages of 1 500 volts.   When correctly
assembled into a counter body (it is important to avoid thermal
shocks at the ceramic/metal bond) they also provide a vacuum-
tight seal at 500° C.   Fig. 11 shows a single coaxial-line design
of counter, using ceramic/metal seals, which is being developed
for high-temperature operation.   The type shown is 0·4 in

diameter, but the design is suitable for other sizes. The anode is completely supported by the ceramic/metal seal and this leads to a simple, robust and rigid design. The extension lead to the counter is a rigid coaxial line with ceramic spacers between the inner and outer conductors.

(4) D.C. OPERATION OF FISSION CHAMBERS

Fission chambers of the type shown in Fig. 6 and 7, coated with uranium 235, have been used as d.c. ionization chambers. The mean current collected in the large chamber of Fig. 7 (with a total U_3O_8 coating of nearly 3 g) is $3 \cdot 4 \times 10^{-7}$ amp for a flux of 10^6 neutrons/cm^2/sec. The current due to the α-activity of the coating is in this case 4×10^{-8} amp, i.e. equivalent to a thermal-neutron flux of about 10^5 neutrons/cm^2/sec. An intrinsic advantage of using uranium 235 as an electrode coating in d.c. ionization chambers instead of boron 10, is that the sensitivity for a given neutron-absorption cross-section may be increased by a factor of about 60 (this being the ratio between the fission-particle energy and the energy of the ionizing particles in the n, α reaction with boron 10). For the measurement of neutron flux in or near the core of small critical assemblies, it is often necessary to limit the total neutron absorption of the detector to a few per cent to avoid undue depression of the flux at the point of measurement. Aves et al.[7] have described a large-area d.c. fission chamber, containing 8 g of uranium 235, used for the measurement of perturbations in neutron flux in oscillator experiments with the Zeus fast reactor; a current of about 1 μA was obtained at a thermal-neutron flux of 10^6 neutron/cm^2/sec.

The lower limit of neutron-flux measurement will be governed by the fluctuations in the measured current due to α-activity; these variations will be due partly to statistical fluctuations and partly to instrumental instability, and will depend on the detailed design of the system. The upper limit of neutron flux will be governed by the ionization-collection efficiency. In d.c. ionization chambers the requirements for collection efficiency are more stringent than in a pulse-type fission counter. In the pulse counter, a recombination loss of 10% or more may only affect the sensitivity at the operating bias level by a fraction of 1% (the effect will depend on the slope of the bias curve). For a d.c. ionization chamber the response is directly proportional to the collection efficiency. By using lower gas pressures than are required for satisfactory pulse-operation, the ionization-collection characteristics can be markedly improved at the expense of some reduction in sensitivity. Table 2 gives some results for a 0·3 cm electrode spacing and 80 cm Hg pressure of argon and a U_3O_8 coating of 0·1 mg/cm^2. The sensitivity is about 20% of the maximum possible (with a sufficiently high gas pressure to dissipate all the fission particle energy within the gas), but with

this reduction in sensitivity it is possible to obtain 99% collection efficiency in fluxes of the order of 10^{11} neutrons/cm^2/sec with a convenient value of polarizing voltage ($\simeq$ 500 volts).

(5) CONCLUSIONS

Fission counters have applications in the measurement fission cross-sections, in the analysis of neutron spectra, and the relative measurements of neutron flux, e.g. in reactor power level instrumentation. The general design criteria applicable all these cases have been summarized in Section 2. The fissi materials used for the first two cases will depend on the particul application and use of the counter. For neutron-flux monitorin uranium 235 is almost invariably used, as this gives the max mum sensitivity and a low α-activity for a given weight of coatin material. For these more general uses it is of interest to compa the fission counter with other neutron detectors, e.g. boro trifluoride proportional counters. A much higher therma neutron sensitivity can be achieved for a given size of count using boron as the detector (the increase may be one to tw orders of magnitude), but the discrimination against a high lev of γ-radiation is much less effective. Thus, with a cathode are of 100 cm^2 a fission counter may still be used in a γ-radiatic dose-rate of 10^5 rads/h; the corresponding limit for a B proportional counter is only a few hundred rads per hour. This property of discrimination against γ-radiation is the ou standing characteristic of the fission counter and often makes the only type of detector suitable for low-level power control small-size reactors, where it is necessary to place the detector or near the core. In future reactor systems these conditions w be associated with high temperatures, and again the fissic counter may be the only detector suitable. A further advantag of the fission counter in instrumentation systems is its lor counting-life. There is at present no evidence of any deteriora tion in characteristics comparable with that observed in BF counters due to dissociation of the counting gas, and fissio counters of the size of the B165 (Fig. 1) have been operated fc some $1 \cdot 5 \times 10^{12}$ counts with no change in counting characteristic

(7) REFERENCES

(1) ROSSI, B. B., and STAUB, H. H.: 'Ionization Chambers an Counters' (McGraw-Hill, 1949).

(2) AVES, R., BARNES, D., and MACKENZIE, R. B.: 'Fissio Chambers for Neutron Detection', *Journal of Nuclea Energy*, 1954, **1**, p. 110.

(3) BAER, W., and SWIFT, O. F.: 'Some Aspects of Fissio Counter Design', *Review of Scientific Instruments*, 195 **23**, No. 1, p. 55.

(4) FACCHINI, U., and MALVICINI, A.: 'Argon-Nitrogen Fillin in Ion Chambers', *Nucleonics*, 1955, **13**, No. 4, p. 36.

(5) COLLI, L., and FACCHINI, U.: 'Drift Velocity of Electron in Argon', *Review of Scientific Instruments*, 1952, 2 No. 1, p. 39.

(6) GILLESPIE, A. B.: 'Signal, Noise and Resolution in Nuclea Amplifiers' (Pergamon Press, 1953).

(7) AVES, R., BAKER, A. R., BARNES, D., and SMITH, R. D 'A Large Fission Chamber Designed for use as a Curren Chamber', A.E.R.E. Memo R/M 119.

(8) ABSON, W., SALMON, P. G., and PYRAH, S.: 'Boron Tr fluoride Proportional Counters' (see next page).

Table 2

D.C. SATURATION CHARACTERISTICS, FOR COUNTER WITH CATHODE COATING OF U_3O_8 (93% URANIUM 235) AT 0·1 MG/CM^2

Area of Coating = 125 cm^2, Electrode Spacing = 0·3 cm
Argon Pressure = 80 cm Hg

Total ionization current	Thermal-neutron flux	Voltage for 99% collection
amp	neutrons/cm^2/sec	volts
$1 \cdot 3 \times 10^{-7}$	$1 \cdot 2 \times 10^8$	150
$1 \cdot 5 \times 10^{-6}$	$1 \cdot 4 \times 10^9$	170
$1 \cdot 6 \times 10^{-5}$	$1 \cdot 5 \times 10^{10}$	250
$8 \cdot 8 \times 10^{-5}$	$8 \cdot 3 \times 10^{10}$	400

Current due to α-activity = $2 \cdot 5 \times 10^{-11}$ amp

BORON TRIFLUORIDE PROPORTIONAL COUNTERS

By W. ABSON, B.Sc., Associate Member, P. G. SALMON and S. PYRAH.

SUMMARY

The design and manufacture of a variety of types of boron-trifluoride proportional counter used by the U.K.A.E.A. are described. One basic design and method of assembly is used for the manufacture of counters of sizes ranging from $\frac{1}{2}$ in diameter and a few inches long to in diameter and several feet long. The technique of manufacture is suitable for large-scale production, and consistently good electrical characteristics can be obtained.

Details are given of the operating characteristics of a wide range of counters, and the effect of circuit parameters on the output pulse amplitude is discussed. Measurements of the γ-response of medium-size counters show that they can be used to discriminate between neutrons and γ-radiation at γ dose-rates of 100–200 rads/h.

The counters can be used for long periods at temperatures up to 100° C without significant deterioration in characteristics. The counting life is found to be a function of counter size and of gas multiplication, and is probably limited by dissociation of BF_3 and the formation of electron-capturing dissociation products. A typical medium-size counter, e.g. 1 in diameter by 6 in long, can be used for 10^{11} counts at a gas-multiplication factor, M, of 10, before significant deterioration of the counting characteristics occurs.

Some applications of BF_3 proportional counters, such as in reactor instrumentation, and the use of large counter arrays in nuclear physics experiments, are discussed briefly.

(1) INTRODUCTION

Boron-trifluoride proportional counters have a wide application in the detection of neutrons in the thermal and epithermal energy region. These applications range from the use of multiple arrays of large counters to provide a high-efficiency detector, e.g. in neutron time-of-flight spectrometers, to the use of very small single counters for flux-distribution measurements, e.g. in experiments on reactor core assemblies and in reactor instrumentation systems.

The ionized track produced by the $^{7}_{3}Li$ and $^{4}_{2}He$ nuclei, which result from the n, α reaction with boron 10, has a length of about $\cdot 5$ cm in BF_3 gas at s.t.p., and the total energy dissipated is about $2\cdot4$ MeV ($2\cdot85$ MeV in 7% of the reactions). Thus, for counters with an effective path length comparable with or greater than this range, all n, α events give rise to a primary ionization considerably greater than that arising from γ interactions, and a good neutron/gamma discrimination is possible. Some degree of gas multiplication is necessary in most practical applications, in order to produce a good signal/amplifier-noise ratio. In order to obtain stable operation as a proportional counter and to obtain a good pulse-height distribution, the BF_3 gas in the counter must be free from electron-capturing impurities; the separation of pure BF_3 gas and the maintenance of counter characteristics over a long period of time present the main difficulties in the design and manufacture of these counters. The main electron-capturing impurities are oxygen and water vapour, and it is particularly important to remove water vapour from both the filling gas and the internal walls of the counter, since BF_3 reacts with water to form HF. This will in turn react

with silica in the glass seal to produce silicon tetrafluoride and water, and the process leads to a gradual deterioration of the counter characteristics.

In previous designs of counters[1] made from brass tubing and turned-brass end-fittings considerable difficulties were encountered in obtaining and maintaining satisfactory vacuum qualities, owing to porosity and age-cracking. Although these defects could be overcome by tinning the outer surfaces with a low-melting-point solder, this process imposed an upper limit of about 100° C on the subsequent outgassing temperatures and long pumping periods were required. A period of ageing with a temporary filling of BF_3, before the final filling of the counter, was found to improve the shelf life of the counters. For high-pressure counters, ageing periods of up to three months have been found necessary.[2]

One of the aims in the new developments to be described was to produce a design of counter with a good shelf-life, which could be manufactured without long periods of ageing. This has been achieved by a design of counter body which can be outgassed at a temperature of 400° C. Large numbers of counters of a variety of sizes are required, and the new design has been adapted to the manufacture of cylindrical counters ranging in size from $\frac{1}{2}$ in diameter and a few inches long to 2 in diameter and several feet long. Neutron absorption in the walls of the counter is kept to a minimum, consistent with strength, by the use of thin-walled copper tubing and thin end-discs. The central anode wire, which is spring-loaded, is supported by a quartz disc at one end of the counter and by a glass/metal seal at the other end, which also provides the connection to the anode. This construction facilitates large-scale production, as simple stamping and rolling operations can be used for the manufacture of piece parts and the location of the quartz insulators in the cathode tubing. A high-grade oxygen-free high-conductivity copper is used to obtain an envelope with good vacuum properties. The internal copper surface is also an excellent getter for any remaining oxygen impurities in the filling gas. R.F. eddy-current heating techniques are used for all brazing operations, forming-gas being used to enable these joints to be made without the use of fluxes and to avoid oxidation on the internal walls of the counter. Experience in the manufacture of these counters has shown that it is possible to obtain consistent and reproducible characteristics with low reject rates and a long shelf-life. It is also found that the performance of counters at elevated temperatures is considerably improved as a result of the higher outgassing temperatures used.

The discriminator-bias/counting-rate characteristics depend on counter dimensions and on gas pressure, the larger counters showing a longer plateau region owing to reduced effects of absorption of the primary particle energy in the cathode wall. Details of the electrical characteristics, counting life and temperature effects are given in Sections 4, 5 and 6. The sensitivities of the counters described cover a wide range, from about $0\cdot06$ counts/sec per unit thermal-neutron flux for short $\frac{1}{2}$ in diameter counters, to 200 counts/sec for long high-pressure counters of 2 in diameter. Some of the applications of these counters are discussed in Section 7.

The authors are at the U.K.A.E.A. Atomic Energy Research Establishment, Harwell.
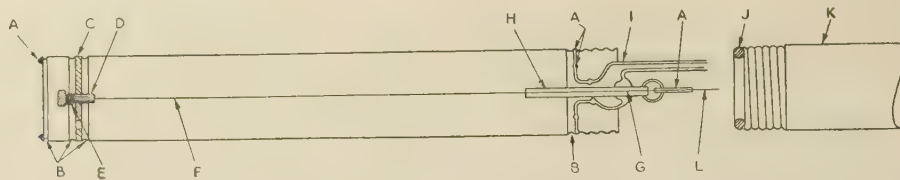
Fig. 1.—General construction of 1 in cathode-diameter BF_3 proportional counter.

Anode diameter = 0·002 in.

Counter body is constructed of high-conductivity oxygen-free copper.

A   Hard-soldered at these four points using silver/copper eutectic solder of melting point 778° C.
B   Swages rolled into body for location of end-caps and for locking quartz disc.
C   Quartz disc.
D   Tube for end termination, spring-loaded.
E   Inconel coil spring.
F   Tungsten or Inconel wire anode.
G   Spot weld to anode wire.
H   Guard-ring electrode.
I   Glass/metal seal.
J   O-ring seal.
K   Screwed end-cap of appropriate length, fitted with optional end termination, provides air-tight joint
      when screwed into counter body.
L   Nickel wire.

(2) DESIGN AND CONSTRUCTION

Fig. 1 shows the general construction of a counter of 1 in diameter; a similar construction is used for those of $\frac{1}{2}$ in and 2 in diameters. The design allows for any of these counters to be made to any desired length, although a preferred range of lengths has been adopted (see Table 3). Similar processes are used for the fabrication and assembly of all types of counter. The main cathode tube is cut to length and a number of the rolling operations, including the thread at the glass-seal end of the counter, are completed. The tube is chemically cleaned and the glass/metal seal end-assembly is brazed in position. The quartz-disc insulator is then fixed in position between grooves formed by a swaging tool. The anode wire assembly is next introduced, prior to the fixing and brazing of the end-disc. All joints are made using copper/silver eutectic solder with eddy-current heating in an atmosphere of forming gas (70% nitrogen and 30% hydrogen). This ensures that all metal parts are kept free from oxidation during heating and joints can be made without the use of flux. An Inconel coil spring [(e) in Fig. 1] is used to hold the anode wire in tension and to allow for relative expansion of the counter body and anode wire during high-temperature outgassing operations before the final gas-filling.

Prior to assembly all metal parts are degreased in acetone, cleaned by an acid dip, and finally washed in hot water and in alcohol. A similar process is used for cleaning the quartz discs.

The anode wire diameters used in the $\frac{1}{2}$ in, 1 in and 2 in diameter counters are 0·001 in, 0·002 in and 0·004 in, respectively. These have been chosen as a practical compromise between efficiency of collection of the primary ionization and gas multiplicaton for a given applied anode voltage. Higher voltages are required with larger diameters or higher gas pressures for a given gas multiplication, and guard-ring seals are fitted in these cases so as to reduce spurious pulses from insulator breakdown. Fig. 1 shows a guard-ring seal for a 1 in diameter counter.

In order to minimize neutron absorption in the counter walls, these are made as thin as possible consistent with strength. The wall thicknesses used for the $\frac{1}{2}$ in, 1 in and 2 in copper tubing are 0·015 in, 0·020 in and 0·025 in,* respectively. In the case of very long 2 in diameter counters it is necessary to support the outside of the tube at 1 ft intervals with close-fitting metal collars to avoid the tube collapsing during outgassing at 400° C. A domed end construction is used in 2 in diameter counters, in place of the flat end-disc shown in Fig. 1, in order to minimize the volume of BF_3 gas in the non-sensitive part of the counter behind the quartz insulating disc.

---

* For a 2 in diameter counter with a copper wall of 0·025 in thickness and filled with $B^{10}F_3$ at 70 cm Hg, the neutron-absorption cross-section of the wall is about 15% of that of B^{10}.

Extension tubes and end adaptors may be screwed in to th rolled thread at the end of the counter tube. The threads ar rolled inwards so that no part of the counter or extension tub need exceed the cathode tube diameter. A rubber washe (J in Fig. 1) may be used to provide an air-tight and water-tigh seal to the adaptor.

(3) GAS PREPARATION AND PURIFICATION: COUNTER FILLING PROCEDURE

Boron trifluoride gas for counter filling is generated from boron-trifluoride/calcium-fluoride complex ($CaF.2BF_3$). Thi complex is a white powder, prepared by heating calcium fluoride in a solution of BF_3 in ether, and distilling off the ether. Details of the method of preparation of the complex and the subsequent preparation of pure BF_3 gas have been described elsewhere.[1,3] Briefly, the method consists in generating BF_3 from the complex by heating at 250° C, having first of all outgassed the complex thoroughly by prolonged heating and pumping at 200° C. As the gas is generated it is passed through cold traps, kept at a temperature of −80° C with a sludge mixture of solid CO_2 and tri-chloro-ethylene, to remove water vapour and other impurities. One of the traps contains sodium fluoride to remove hydrofluoric acid. The BF_3 gas is then frozen in a further trap surrounded by liquid nitrogen. After a suitable quantity of solid BF_3 has been produced, the liquid-nitrogen trap is closed off from the generating system. This trap is then allowed to warm up and the gas is transferred to glass storage vessels which have previously been outgassed. The storage vessels are connected to a multi-way glass manifold to which the counters are connected by means of the glass filling stems on the glass/metal seals. The counters are continuously evacuated and baked at a temperature of 400° C for three hours, after which they are filled with BF_3 at about atmospheric pressure from a subsidiary storage vessel. This gas is not the final filling gas but is used as a temporary filling for a few hours to 'soak' the counter bodies and to react with any remaining impurities in the counter envelope. This gas is removed and the counters are pumped out to a hard vacuum at room temperature. BF_3 gas from the storage vessels is then frozen into a trap connected by a mercury cut-off valve to the filling manifold. The frozen BF_3 is allowed to warm up until the counters are filled to a pressure slightly less than atmospheric, and the trap is then cooled again to reduce the pressure by a factor of 2 or 3. This process is carried out quickly at least three times, in order to enable all the gas to come into contact with the cold trap but without completely freezing it again; remaining impurities with a higher freezing point may thus be removed. The gas pressure is finally allowed to rise to the

quired filling pressure and the mercury cut-off valve is closed. ach counter is then sealed and removed from the filling manifold heating and drawing off the glass filling stem.

(4) ELECTRICAL CHARACTERISTICS

The kinetic energy of the ionizing particles resulting from the α reaction with boron 10 is 2·4 MeV (2·85 MeV in 7% of e reactions). The maximum pulse-amplitude available from an ent in which the whole of the particle energy is dissipated in e gas is given approximately by

$$\frac{M}{C} \times 10^{-2} \text{ volts*}$$

here M is the gas multiplication and C is the total input capaci- nce in picofarads. This maximum pulse size will only be otained, however, if the effective differentiating time-constant in e amplifier is large compared with the total positive-ion transit

amplifier unit, although it is often advantageous to house this circuit in a sealed and desiccated case to reduce spurious pulses due to breakdown across insulators caused by dust or condensed moisture. Similarly the anode lead-through seal of the counter should also be kept free from dust and condensed moisture. Counters operating at voltages in excess of 2 kV are normally fitted with a guard-ring seal to reduce spurious breakdown pulses, the guard electrode being connected to a low-impedance point in the e.h.t. filter circuit. The value of the anode-load resistor R is chosen so that the time-constant RC of the input circuit is large compared with the differentiating time-constant used in the later stages of the amplifier. The amplifier is usually composed of two sections, a pre-amplifier unit with a fixed gain (in the range 10–100), which is placed as near as possible or convenient to the counter or counters so as to minimize the input capacitance C, and a main amplifier section containing power supplies and further stages of amplification (gain adjustable from 10^2 to 10^4) and pulse-shaping circuits (differentiating and integrating circuits).
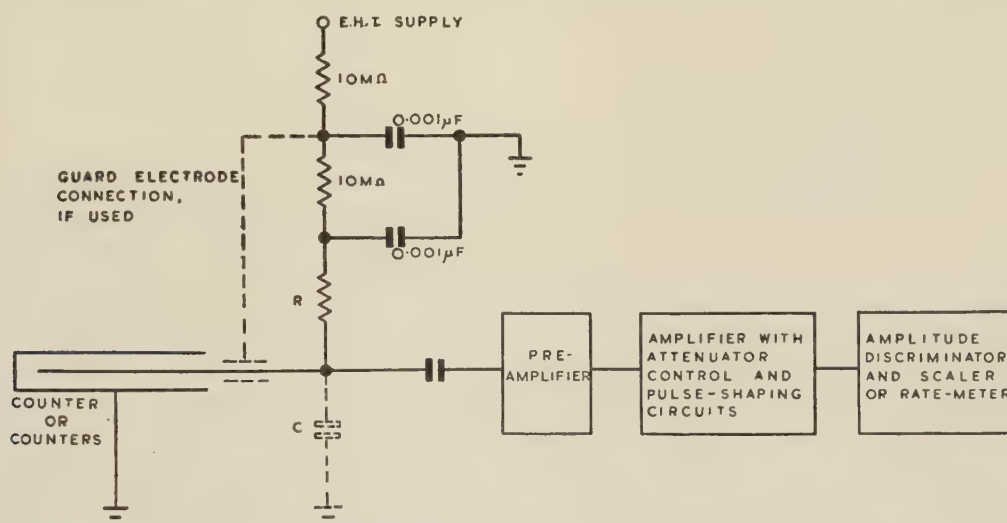
Fig. 2.—Schematic of counter, amplifier and pulse-counting circuit.

imes in the counter (i.e. of the order of 100 microsec). The oltage induced on the anode wire by electrons from the primary ionization or from gas-multiplication processes is negligible compared with that due to the movement of the positive ions ormed near the anode. The voltage pulse rises steeply at first and then more slowly as the positive ions move towards the cathode; the pulse attains about one-half of the maximum amplitude in a time of the order of a microsecond, but the total transit time of the positive ions to the cathode is of the order of 100 microsec. It is usual to make use of only the steeply rising part of the pulse by differentiating the signal in the later stages of the pulse amplifier, and the effective pulse amplitude is reduced by a factor of 2 or more. The energy dissipated in the gas by the primary ionizing particles may be less than the maximum (2·4 MeV) when the range of the particles is comparable with the counter diameter, and this further reduces the pulse amplitude. Taking these two effects into account, the input sensitivity required in the amplifier is usually in the range $(M/C) \times 10^{-3}$ to $(M/C) \times 10^{-4}$ volts.

Fig. 2 shows a schematic of a typical arrangement of pulse amplifier and pulse-counting circuit used in fixed installations with BF$_3$ proportional counters. The e.h.t. filter circuit and the counter anode-load resistor R may be mounted in the pre-

(4.1) Pulse Amplitude Distribution: Integral and Differential Bias Characteristics

Counting-rate/discriminator-bias characteristics for 1 in dia-meter counters at three gas pressures (20, 40 and 70 cm Hg) are shown in Fig. 3. Half-inch diameter counters with 40 cm Hg pressure fillings have a bias characteristic similar to curve (c) in Fig. 3 for a 20 cm Hg 1 in diameter counter. The maximum lengths of the ionized tracks in the BF$_3$ gas at 20, 40 and 70 cm Hg pressure are approximately 2·0, 1·0 and 0·6 cm, respectively. The improved integral bias curves obtained at higher gas pres-sures are attributed to a reduced wall effect, i.e. the proportion of particles dissipating a large fraction of their energy in the walls is reduced. This is further demonstrated by the differential bias characteristic for 1 in and 2 in diameter counters filled at 40 cm Hg, shown in Fig. 4. In each case the whole counter was uniformly irradiated. In the larger and higher-pressure counters it is particularly important to avoid traces of electro-negative impuri-ties in the gas, as the probability of electron capture is increased owing to the longer paths and collection times involved. If an electron is captured by an impurity molecule to form a heavy negative ion it is no longer effective in producing gas multiplica-tion. Ionizing tracks produced near the cathode will lose a larger proportion of their electrons than tracks formed near the anode wire, and the pulse-height distribution will be broadened. The precautions mentioned in Sections 2 and 3, during gas

* The maximum charge per pulse is $MW_p e/W_i$, where W_p is the particle energy in electron volts, W_i is the energy required per ion pair ($\simeq 33$ eV in BF$_3$) and e is the electronic charge.
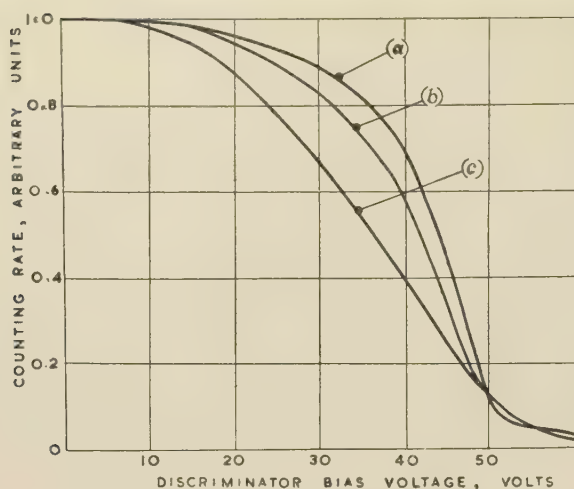
Fig. 3.—Integral bias curves for 1 in diameter BF_3 counters with 0·002 in diameter anode wires.

(a) BF_3 Pressure 70 cm Hg; anode voltage 2 400 volts; $M = 40$.
(b) BF_3 Pressure 40 cm Hg; anode voltage 1 800 volts; $M = 40$.
(c) BF_3 Pressure 20 cm Hg; anode voltage 1 200 volts; $M = 40$.

Amplifier setting: Gain = 2×10^4.
Differentiating time-constant = 0·8 sec.
Integrating time-constant = 0·08 sec.
Total input capacitance = 120 pF.
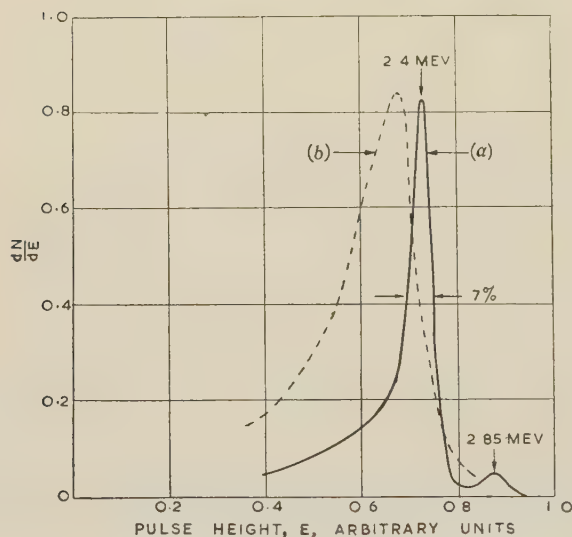
Fig. 4.—Differential-bias curves.

(a) 2 in diameter counter ⎱ BF_3 pressure 40 cm Hg.
(b) 1 in diameter counter ⎰

purification and out-gassing of filling plant and counter bodies, have been found to be essential in order to obtain and maintain counting characteristics of the standard shown in Figs. 3 and 4.

The $_2He^4$ and $_3Li^7$ particles, resulting from the neutron reaction with B^{10}, are emitted in opposite directions and at least one of them will produce an ionized track in the BF_3 gas. Except in the case of very small or very-low-pressure counters, the energy dissipated will be sufficient to produce an observable pulse, and hence a 100% detection efficiency* will be obtainable. In all the cases shown the slope of the counting-rate/discriminator-bias characteristic is virtually zero at a bias setting equivalent to 0·1 of the maximum pulse height, and 100% detection efficiency is obtained below this bias level. It may be necessary to operate

* Detection efficiency is defined here as the ratio of the number of observed counts to number of n, α events within the sensitive volume of the counter.

counters at higher bias levels, in order to avoid spurious count[ing] from electrical interference or perhaps from the 'pile-up' due t[o] high γ-radiation levels. Long-term stability can still be achieve[d] without stringent requirements on overall gain stability at effec[ti]ve bias levels of 0·2–0·4 of the maximum pulse height. Table [1]

Table 1

COUNT-RATE AND DIFFERENTIAL COUNT-RATE AT EFFECTIVE BIAS
LEVELS 0·2 AND 0·4 OF THE MAXIMUM PULSE HEIGHT

Type of counter		Bias level (0·2 of maximum pulse height)		Bias level (0·4 of maximum pulse height)	
Diameter	Gas pressure	Detection efficiency	Change in count-rate for 10% bias change	Detection efficiency	Change in count-rate for 10% bias change
in	cm Hg	%	%	%	%
$\frac{1}{2}$	40	>98	<1	>85	<4
1	20	>98	<1	>85	<4
1	40	>99	<0·4	>93	<2
1	70	>99	<0·4	>95	<1·5
2	70	>99	<0·4	>95	<1·5

gives information on the detection efficiency observed at these higher bias levels, and also the slope of the bias characteristic expressed as the percentage change in count-rate for a 10% change in overall gain or effective bias level. The figures quoted are representative of results obtained with large numbers of counters.

The bias characteristics shown in Fig. 3 refer to counters operating at $M = 40$, amplifier differentiating and integrating time-constant of 0·8 microsec and 0·08 microsec, and a total input capacitance of 120 pF. These conditions are typical of many practical applications, and the amplifier input sensitivity required, for an operating point on the bias curve equivalent to 0·2 of the maximum pulse height, is about $4·8 \times 10^{-4}$ volt. The equivalent input charge sensitivity is $5·7 \times 10^{-14}$ coulomb, or $3·6 \times 10^5$ ion pairs. The input charge sensitivity of a counter-amplifier system may be measured by injecting a voltage step-function, V, into the input via a series capacitance C_s. When C_s is very small compared with the total input capacitance C, the injected charge is $C_s V$. The voltage generator may then be calibrated in terms of coulombs or ion pairs. If $C_s = 0·16$ pF a voltage step $V = 1$ volt corresponds to 10^6 ion pairs.

(4.2) Gas Multiplication

The gas multiplication is a function of the counter anode and cathode diameters, the gas pressure and the anode voltage. The multiplication process starts when electrons near the anode wire can gain sufficient energy from the electric field, between collisions, to produce ionization. For cylindrical counter geometry the voltage V_c at which gas multiplication starts is given by

$$V_c = X_c \frac{pa}{\log (b/a)} \quad \cdots \quad (1)$$

where X_c is a property of the gas, equal to E_c/p, where E_c is the critical field strength for gas multiplication, p is the gas pressure, and b and a are the cathode and anode radii. An analysis by Rose and Korff[4] gives the following general expression for the gas multiplication M at an applied voltage V:

$$M = \exp \left\{ \left[\frac{fpaV}{\log (b/a)} \right]^{1/2} \left[\left(\frac{V}{V_c} \right)^{1/2} - 1 \right] \right\} \quad \cdots \quad (2)$$

where f is a constant depending on the gas. As V is increased above the value V_c, M increases slowly at first and then more rapidly to a region where $\log M$ is almost linear with V.

Fig. 5 shows the gas-multiplication/anode-voltage characteristics for a range of BF$_3$ counters of various sizes and gas
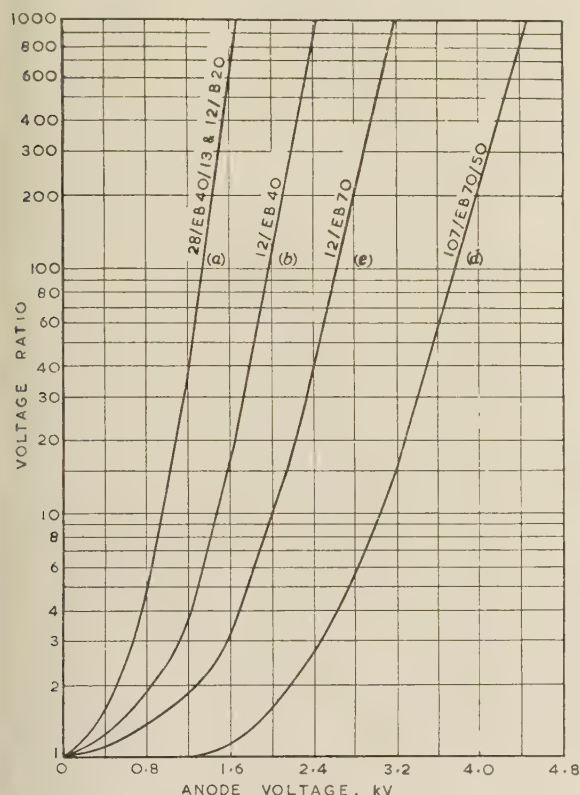
Fig. 5.—Gas multiplication, M, as a function of anode voltage.

(a) 1 in cathode diameter, 0·002 in anode diameter, BF$_3$ pressure 20 cm Hg; and ½ in cathode diameter, 0·001 in anode diameter, BF$_3$ pressure 40 cm Hg.
(b) 1 in cathode diameter, 0·002 in anode diameter, BF$_3$ pressure 40 cm Hg.
(c) 1 in cathode diameter, 0·002 in anode diameter, BF$_3$ pressure 70 cm Hg.
(d) 2 in cathode diameter, 0·004 in anode diameter, BF$_3$ pressure 70 cm Hg.

pressures. The variation of gas multiplication with anode voltage was obtained by measuring the change in the bias setting required to give a counting rate of 20% of the plateau counting rate. The relative values of gas multiplication for the various types of counters are accurate, but the absolute value of M is only approximate and was obtained by adjusting the ordinate scale of Fig. 5 so that M approaches unity asymptotically at low voltages. The increased voltage required to achieve a given multiplication in the larger-diameter counters is due partly to the use of larger-diameter anode wires. Although the field strength near the anode is increased by the use of thinner anode wires and the gas multiplication hence increased, the field strength near the cathode is reduced and some compromise is necessary to avoid recombination effects in the initial ionizing tracks, owing to an inadequate field strength near the cathode. The compromise adopted in the counters described is to use 0·001 in, 0·002 in and 0·004 in diameter anode wires in counters with ½ in, 1 in and 2 in diameter cathodes, respectively (i.e. $b/a = 500$ for all counters). It will be seen from eqn. (2) that, if b/a is constant and the product pa is constant, the multiplication/anode-voltage relation is identical for counters of any diameter. This is demonstrated in curve (a) of Fig. 5, which applies to ½ in diameter counters at 40 cm Hg and to 1 in diameter counters at 20 cm Hg pressure.

For $M > 10$ the experimental curves can be represented by the relation

$$\log_{10} M = \frac{(V - V_k)}{V_{10}} \quad \ldots \ldots \quad (3)$$

where V_{10} is the slope of the straight-line portion of the curves in Fig. 5, and V_k is the intercept when extrapolated to $M = 1$. Both are functions of pa when b/a is constant. V_{10} is the voltage change required to increase M by a factor of 10. The results given in Fig. 5 cover a range of pa from 0·05 to 0·35, p being in centimetres of mercury and a in centimetres; numerical values for V_{10} and V_k are given in Table 2.

Table 2

VALUES OF V_k AND V_{10} IN EQN. (3) FOR BF$_3$ COUNTERS IN WHICH $b/a = 500$

pa	V_k	V_{10}
cm Hg × cm	volts	volts
0·05	550	400
0·10	960	510
0·17	1 400	570
0·35	2 330	715

For most practical purposes, values of V_k and V_{10} for intermediate values of pa can be obtained by linear interpolation in Table 2.

(4.3) Effect of Differentiating Time-Constant on Pulse Height

The amplifier time-constants used with proportional counters are usually short compared with the total transit time of the positive ions, and the output pulse shape is determined by the rise time of the initial fast part of the counter pulse and by the value of the amplifier time-constants. If the primary ionized track is parallel to the anode wire, the pulse will occur after a delay equal to the transit time of the electrons to the wire, and the pulse-rise time will be governed by the subsequent movement of positive ions away from the anode. For initial tracks oriented in other directions, there will be a spread in the transit time of the electrons, which will affect the initial shape of the voltage pulse at the anode, since the spread in electron transit time is of the order of a microsecond.

Fig. 6 shows the change in the average pulse amplitude as the
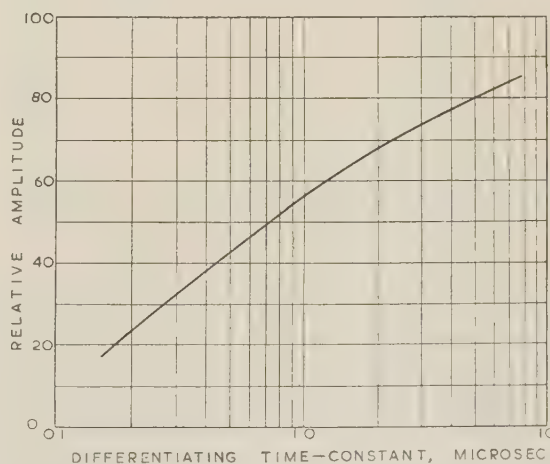
Fig. 6.—Effect of differentiating time-constant on pulse amplitude.

Type 12EB40 counter. 1 in cathode diameter, 0·002 in anode diameter. Integrating time-constant = 0·08 microsec.

amplifier differentiating time-constant, T_1, is changed from 8·0 microsec to 0·16 microsec for a 1 in diameter counter with a BF$_3$ pressure of 40 cm Hg. The counter was operated at $M = 40$ and the amplifier integrating time-constant, T_2, was 0·08 microsec. Gillespie[5] has suggested that in proportional counters in which the initial ionizing tracks are distributed at random the majority of pulses have an initial linear rise. With this assumption, it can be shown that the results of Fig. 6 are consistent with an initial rise-time, T, of 1 microsec. When the amplifier time-constants T_1 and T_2 are small compared with T, the resolving time of the system will be approximately equal to T. The counting losses at high counting rates will then be given by $100\bar{n}T$ per cent, where $\bar{n}$ is the mean counting rate. Thus for $T = 1$ microsec, 10% counting losses will occur at a mean counting rate of 10^5 counts/sec. For a counter filled at higher pressure and operated at the same voltage, the rise time T will be increased in proportion to the pressure increase; if the counter is used at the same gas multiplication, however, this increase will be partially offset by the higher anode voltage required.

(4.4) Response to γ-Radiation

The magnitude of the individual pulses due to γ-ray interactions in the counter walls and gas is determined by the counter dimensions and the gas pressure. For a 1 in diameter counter filled with BF$_3$ to a pressure of 40 cm Hg the average pulse height due to 1 MeV γ-radiation is of the order of one-hundredth of the maximum 'neutron' pulse height. When the mean interval between γ pulses becomes less than the resolving time of the counter and associated circuits, pulse build-up will occur and counts due to γ-radiation will be registered at higher discriminator-bias levels.

The effect of cobalt 60 γ-radiation, at dose-rates from 25 to 250 rads/h, on a type 12EB40 counter (6 in × 1 in and 40 cm Hg pressure of BF$_3$) is shown in Fig. 7. The amplifier time-constants
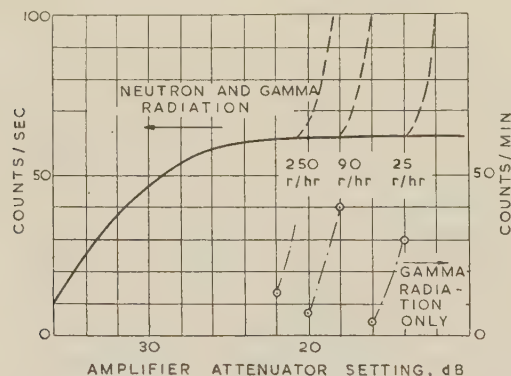
Fig. 7.—Effect of high γ-radiation levels.

Type 12EB40 counter with cobalt 60 source.
Counter anode voltage = 1 800 volts ($M = 40$).
Amplifier settings: Differentiating time-constant = 0·16 microsec.
         Integrating time-constant = 0·16 microsec.
1 in cathode diameter, 0·002 in anode diameter, BF$_3$ pressure
      40 cm Hg.

T_1 and T_2 were 0·16 microsec, and the resolving time of the system was therefore approximately equal to the rise time of the counter pulse (1 microsec). The full-line curve shows the counting-rate/amplifier-gain characteristic for a neutron source in the absence of a high level of γ-radiation, and the broken lines indicate the abrupt increase in counting rate due to γ-radiation levels of 25, 90 and 250 rads/h. The chain-dotted curves were obtained from measurements with the neutron source removed and show the γ-response characteristics at low counting rates.

An approximate theory of the effect may be developed by considering the fluctuations in the number of γ events occurring within the resolving time T_r of the system. If N is the number of ionizing events per second, the standard deviation in the number occurring within a time T_r will be $\sqrt{(NT_r)}$; if p_z is the probability that the deviation is $z\sqrt{(NT_r)}$, a counting rate of p_z/T_r per second will be observed at a discriminator bias level of $Bz\sqrt{(NT_r)}$ where B is the bias level of a single γ-pulse. The value of N will be proportional to the γ dose rate and to the cathode area of the counter (or counters, if several are used in parallel), whilst the average value of B will be a function of counter dimensions and gas pressure. The results shown in Fig. 7, for the dose rates 25 rads/h to 250 rads/h, show a smaller variation in the critical bias level for onset of γ-response with γ-radiation level than is indicated by the simple theory; a change of 10 : 1 in γ dose rate is accompanied by a change of 7 dB in critical bias level, compared with a factor $\sqrt{10}$, or 10 dB, given by the theory. This discrepancy is probably due to the spread in the size of the individual γ-pulses. The formula is useful, however, in indicating the magnitude of the γ-response for counters of different size and gas pressure. For a given size of counter an increase in pressure will produce a proportionate increase in B, and T_r will also be increased. Thus, for a type 12EB70 counter (6 in × 1 in and 70 cm Hg gas pressure) the critical bias level for onset of γ-response will be higher, by a factor of about 2, than the levels shown in Fig. 7 for a counter filled to 40 cm Hg pressure.

It is often convenient to have the shortest time-constant clipping circuit at an intermediate stage in the amplifier. In this case it is necessary to ensure that the input time-constant is sufficiently small to avoid overloading of the earlier stages of the amplifier by the signal fluctuations arising from the γ-radiation. The results shown in Fig. 7 were obtained using an input time-constant of about 5 microsec; the head amplifier gain was 100 and the clipping circuit followed this amplifier.

With high γ-radiation levels and high gas-multiplication the positive-ion space charge in the counter may be sufficient to alter the field distribution and cause a reduction in gas multiplication. The effect is observable in a type 12EB40 counter at a dose rate of 250 rads/h and a gas multiplication of 40, the neutron count-rate characteristic being shifted by about 0·5 dB. (This effect is not shown in Fig. 7.)

Spurious counts may also arise in high γ-radiation fields due to the collection of charges and their subsequent redistribution on the anode insulating supports. The effect is most marked where there are insulated surfaces with only a small potential gradient along the surface, e.g. a glass sleeve terminating the anode wire at the quartz-disc end of the counter. The results shown in Fig. 7 were obtained with a glass sleeve of about $\frac{3}{8}$ in overall length and are similar to those obtained when the glass sleeve is replaced by a metal tube. A number of counters have been made with longer glass sleeves, and these give rise to spurious counts, although the effect is important only at low counting rates (a few per second) and for counters used at low gas-multiplication ($M \simeq 10$).

(5) COUNTING-LIFE

A gradual deterioration of counting characteristics has been observed after continuous operation for long periods at high counting rates; this effect is more marked for counters operated at high gas-multiplication. Results obtained with 6 in × 1 in counters filled to a pressure of 40 cm Hg and operated with gas-multiplication factors of 8, 40 and 400 are shown in Fig. 8. The counting rate used in the course of these experiments was about 2×10^4 counts/sec, and the comparatively minor deterioration in curve (b) ($M = 8$ and 10^{11} counts operation) corresponds to continuous operation at this level for a period of eight weeks. At the higher values of M (40 and 400) the change in pulse height
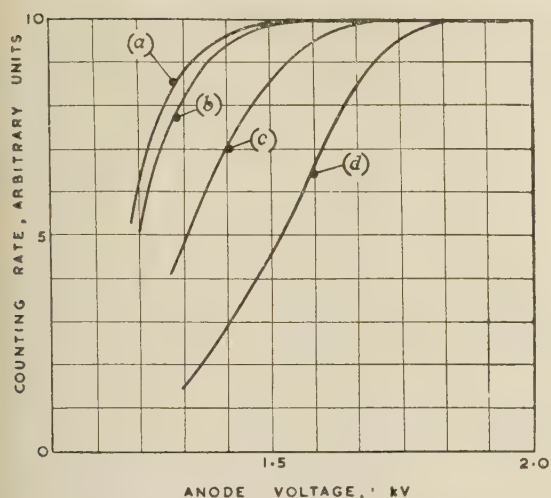
Fig. 8.—Effect of large number of counts on standard 1 in counters filled at 40 cm Hg.

(a) Initial h.t. counting-rate characteristic.
(b) After 10^{11} counts for counter operated at 1 400 volts ($M = 8$).
(c) After 2×10^{10} counts for counter operated at 1 800 volts ($M = 40$).
(d) After $1 \cdot 3 \times 10^{10}$ counts for counter operated at 2 250 volts ($M = 400$).

and pulse-height distribution becomes observable after about 10^{10} and 10^9 counts operation.

The change in the characteristics is similar to that produced by an electron-capturing impurity in the gas. Soberman *et al.*[6] have reported results for BF_3 counters operating at very high gas-multiplications and counting γ-pulses. They have attributed the deterioration in characteristics to the production of fluorine and BF_2 from dissociation of BF_3. A mechanism of this type is consistent with the observed dependence of the effect on gas multiplication.

It is found that the characteristics of counters damaged in this way may be recovered by heating the counter for a period of three to four hours at about 180° C. The recovery process is gradual, and the effect of a four-hour heating period is shown in Fig. 9. The subsequent performance of the counter after 8×10^{10} counts operation at $M = 40$ is also shown. It will be seen, by comparison with curve (a) of Fig. 9, that the deterioration effect is much less than that obtained originally. The recovery, after

Fig. 9.—Recovery of characteristics after heat treatment.
1 in diameter counter filled at 40 cm Hg.

(a) Deterioration after 6×10^{10} counts operation at 1 800 volts ($M = 40$).
(b) Recovery after heating for 4 hours at 200° C.
(c) Subsequent performance after a further 8×10^{10} counts operation at 1 800 volts.

heat treatment, is no doubt due to the absorption of the impurities in the copper walls. Copper acts as a getter for fluorine, but it would be expected that the amounts produced would be absorbed in the copper without extended heat treatment. The reduction in the deterioration effect after initial deterioration and subsequent recovery with heat treatment is not fully understood.

(6) OPERATION AT ELEVATED TEMPERATURES

The construction of the counters described allows for subsequent uniform heating to 400° C without physical damage to the counter body. At such temperatures, however, the counting characteristics deteriorate quickly, owing to accumulation of electron-capturing impurities given off from the counter walls. This effect is negligible at 100° C for counters which have been previously outgassed at 400° C, and counters have been operated continuously at 100° C for a period of three months without appreciable change in characteristics. At 150° C a change in characteristics may be observed after a period of a few weeks.

The high-temperature performance of BF_3 proportional counters depends on the previous outgassing treatment during manufacture. In an earlier design of counter made at the Atomic Energy Research Establishment,[1] where the maximum outgassing temperature was 120° C, a marked deterioration in characteristics occurred after a few hours' operation at 80° C. Similar results showing a deterioration occurring at 70° C have been reported by Schultz and Connor.[7]

(7) SOME APPLICATIONS OF BF_3 PROPORTIONAL COUNTERS

Table 3 lists some standard lengths and gas fillings, for $\frac{1}{2}$ in, 1 in and 2 in diameter counters, used at the Atomic Energy

Table 3

STANDARD COUNTERS USED AT THE ATOMIC ENERGY RESEARCH ESTABLISHMENT

Type	Diameter	Approximate overall length	Anode voltage for $M = 40$	Approximate sensitivity*
	in	in	volts	counts/sec per unit thermal-neutron flux
5B40/13	0·5	4·5	1 200	0·06
5EB40/13	0·5	4·5	1 200	0·3
28EB40/13	0·5	13·5	1 200	1·7
12B20	1	7	1 200	0·3
12EB40	1	7	1 800	3·0
12EB70/G†	1	7	2 400	5·0
31EB40	1	14·5	1 800	7·6
31EB70/G	1	14·7	2 400	13·5
15EB70/50G	2	9	3 500	27·5
40EB70/50G	2	19	3 500	73
107EB70/50G	2	45	3 500	196
84EB45/50G	2	36	2 600	98

* The sensitivity is calculated from the total cross-section of the boron 10 within the sensitive volume.
† A final letter G in a type number denotes the use of a guard-ring electrode in the glass/metal seal.

Research Establishment. The number preceding the letters B or EB in the type designation denotes the active length in centimetres. B denotes natural boron and EB denotes boron enriched in the boron 10 isotope to about 95%. The next number denotes the gas pressure. For counters other than those of 1 in diameter a final number is added which gives the diameter in millimetres.

The $\frac{1}{2}$ in diameter counter has been developed mainly to fulfil
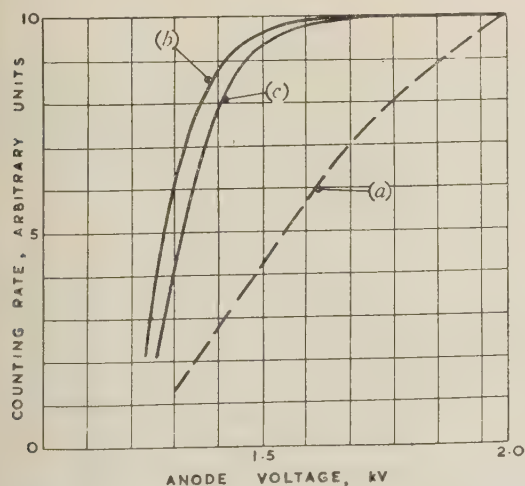
requirements for flux-distribution measurements, e.g. in criticality experiments with reactor core assemblies. The counter is usually attached to a rigid extension lead and the whole probe assembly need not exceed $\frac{1}{2}$ in in diameter. A water-tight joint can be obtained by the use of a rubber washer between the counter and the extension lead, for measurements in assemblies under water.

One-inch diameter counters are used in a wide variety of applications, e.g. in loading experiments in reactors, reactor control instrumentation, and as detectors in the measurement of neutron cross-sections in the thermal and epithermal energy regions. In the latter type of experiments large arrays of counters are used to obtain a high-efficiency detector. A convenient method of assembling a closely packed bunch of counters is to screw them to adaptors which have been brazed on to a flat metal plate. The head amplifier and e.h.t. filter unit can then be mounted on the other side of this plate and enclosed within a sealed cover. This construction ensures that all insulating surfaces, including the glass/metal seal on each counter, are kept clean and dry, and that the total input capacitance is kept to a minimum. An array of this type used at the A.E.R.E. consists of 36 type 31EB70 counters connected in parallel and operating from a common e.h.t. supply.

The largest assembly of BF$_3$ counters used so far at the A.E.R.E. comprises 240 counters, each of 2 in diameter and 4 ft long. The counters are connected in groups of 10 or so to head-amplifier units with cathode-follower outputs; these outputs are fed into a main amplifier and a multi-channel pulse-
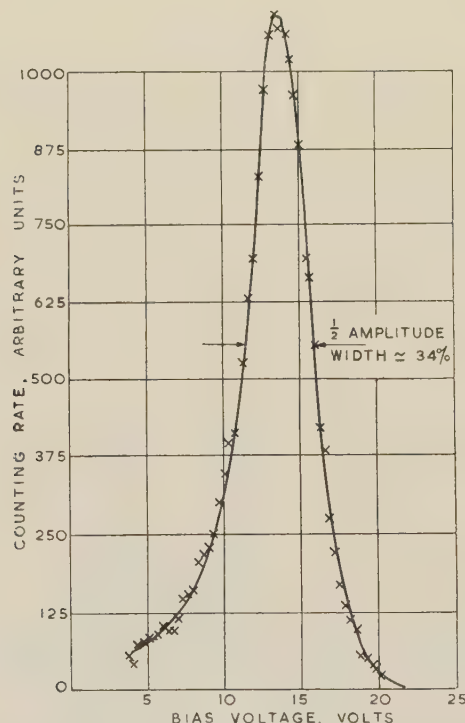
Fig. 10.—Differential-bias curve of 76-counter array.
Each counter of 2 in diameter and 4 ft length (type 107EB70/50G).

amplitude analyser. Fig. 10 shows the differential-bias curve for 76 of these counters connected in 10 groups to 10 head-amplifier units.

In neutron time-of-flight spectrometer experiments the neutron-energy resolution obtainable depends on the delay time and spread in delay time of the detector. In a BF$_3$ proportional counter the time delay between an n, α reaction and the rise

of the counter output pulse will depend on the position and orientation of the initial ionizing track and the size, gas pressure and operating voltage of the counter. Nicholson[8] has measured the time spread or jitter in the pulses for 1 in and 2 in diameter counters of the type described here. The distribution of delay times is rectangular; the total width for a 1 in diameter counter filled at 40 cm Hg and operated at 2 400 volts is 0·6 microsec; for a 2 in diameter counter filled at 70 cm Hg and operated at 4 000 volts the width is 2·75 microsec.

Pulse-type counters may be required in reactor control instrumentation systems for neutron-flux measurement at sub-critical levels and during start-up of the reactor, when the ratio between neutron and γ-radiation fluxes is too low for the use of mean-current ionization chambers. The limiting values of γ-radiation dose rate at which BF$_3$ proportional counters may be used, and discrimination obtained between neutron pulses and the effects of γ-radiations, have been discussed in Section 4.4. When the reactor power level increases beyond the range of the counter, it is necessary to retract the counter to a position of reduced neutron flux to avoid excessive activation of the counter body, so that it can be used again at low levels when the reactor is shut down. It is also necessary to switch off the h.t. anode supply to avoid the deterioration effects which result from a large number of counter discharges (see Section 5).

The effect of a high β-γ activity induced in the materials of the counter is similar to that of a high γ-radiation dose rate. Types 12B20 and 12EB40 counters, irradiated for 10 days in a thermal-neutron flux of 3×10^9 neutrons/cm^2-sec (i.e. integrated flux of $2 \cdot 5 \times 10^{15}$ neutrons/cm^2) with zero anode voltage, show no deterioration in their characteristics. The 12·8-hour β-γ activity due to copper 64 immediately after removal from a flux of 3×10^9 neutrons/cm^2-sec is, however, too high to enable neutron pulses to be distinguished. After a decay period of about 7 half-lives (i.e. a factor 128 reduction in activity), good discrimination can be obtained between neutron pulses and the effect of the copper 64 activity. Counters of this size should therefore be retracted to a position where the thermal-neutron flux is less than about 2×10^7 neutrons/cm^2-sec at full-power operation. The saturation activity will then be sufficiently low to avoid difficulty in discriminating between neutron pulses and β-γ 'pile-up'. At this flux level the exposure period for an integrated flux of $2 \cdot 5 \times 10^{15}$ neutrons/cm^2 will be approximately three years.

(8) CONCLUSIONS

Counters of the type described here have been produced in large quantities and the technique of manufacture gives consistent and reproducible results. The design is adaptable to the wide range of counter sizes and sensitivities required in instrumentation systems and in experimental work for the measurement of neutron flux in the thermal and epithermal energy range.

(9) ACKNOWLEDGMENTS

(10) REFERENCES

(1) SHARPE, J., and SALMON, P. G.: 'The Manufacture of Boron Trifluoride Proportional Counters', A.E.R.E. Report No. EL/R1239.
(2) FOWLER, I. L., and TUNNICLIFFE, P. R.: 'Boron Trifluoride

Proportional Counters', *Review of Scientific Instruments*, 1950, **21**, p. 734.

HUDSWELL, F., NAIRN, J. S., and WILKINSON, K. L.: A.E.R.E. Report No. C/R 651.

ROSE, M. E., and KORFF, S. A.: 'An Investigation of the Properties of Proportional Counters', *Physical Review*, 1941, **59**, p. 850.

GILLESPIE, A. B.: 'Signal, Noise and Resolution in Nuclear Amplifiers' (Pergamon Press, 1953).

(6) SOBERMAN, R. K., KORFF, S. A., FRIEDLAND, S. S., and KATZENSTEIN, H. S.: 'Deterioration of Boron Trifluoride Counters due to High Counting Rates', *Review of Scientific Instruments*, 1953, **24**, p. 1058.

(7) SCHULTZ, M. A., and CONNOR, J. C.: 'Reactor Power Calibration', *Nucleonics*, **12**, No. 2, p. 8.

(8) NICHOLSON, K. P.: 'Delays in BF$_3$ Proportional Counters', A.E.R.E. Report No. N/R1639.

DISCUSSION ON THE ABOVE FOUR PAPERS BEFORE THE MEETING OF THE MEASUREMENT AND CONTROL SECTION HELD IN CONJUNCTION WITH THE BRITISH NUCLEAR ENERGY CONFERENCE, 4TH FEBRUARY, 1958

Dr. D. Taylor: The paper by Dr. Howlett gives an admirable survey of the present position regarding digital computers and their use in reactor-design calculations. He mentions the developments in mathematical techniques which have taken place in the last few years, which have tended to reduce the work which must be done by the computer, and the question therefore arises whether larger and faster computers are really required, or whether improvements in the mathematical methods would not pay dividends more quickly. Is the British equivalent of Stretch required if we are to maintain our present position in the nuclear-power world, or are advances in the mathematical techniques for handling design calculations likely?

The paper by Messrs. Bowen and Masters deals with the use of analogue computers for carrying out certain reactor calculations. Some experimental confirmation of the results has been obtained from the work done on Bepo. It would be of interest to know if any further experimental work has been carried out bearing on these calculations.

The first paper by Messrs. Abson, Salmon and Pyrah is of special interest. It is not very long ago that the making of a BF$_3$ proportional counter was regarded as something of a mystery. As a result of the work now reported, proportional counters of this kind are made in relatively large numbers with very reproducible characteristics. Fig. 8 apparently illustrates a method of making old counters new, but it also means that such counters cannot be used at high temperatures. The difficulties caused by the production of free fluorine at elevated temperatures could be avoided by introducing boron as the element and as deposition on one or both electrodes. Has this method been explored, and does it introduce any further difficulties? Another point of some importance is the upper limit of neutron flux in which a BF$_3$ counter can remain. This is presumably set by induced activities, and so it may be possible to do better by using something other than copper. Has this been explored? If so, information would be of interest.

The last paper deals with fission counters, and in it the authors refer to the design of a high-temperature fission counter. Presumably the counter mentioned has now been completed. It would be of interest to know whether it has been a success.

It may be noted that the BF$_3$ proportional counter tends to be used when one wants a high sensitivity and when the γ-background is low. When the γ-background is high the answer seems to be to use a fission counter and endure the reduced sensitivity. Is there a way out of this, i.e. is it possible to improve the sensitivity of the fission counter?

Mr. F. R. Farmer: From Dr. Howlett's paper I got the impression that computing machinery is likely to become more and more complicated, while in that by Messrs. Bowen and Masters saw an attempt to simplify these problems to produce some scale whereby reactors of new types can be compared. In the long run we would agree that the only way to find out how a reactor behaves is to build one; but even when it is built, experiments carried out on the reactor are difficult to interpret unless theory lies behind them and unless also we have some means of using this theory in some quick and reproducible method. We can and do build analogue machines which enable us to evaluate particular patterns of reactor. But the response of these machines is only the response which we as designers put into them. This seems to me to be going round in a circle.

Yet looking at other aspects of man's achievement, I am sure that very few engineers know very much about the properties of all the metals they use. But series of equations have been produced which enable them to build structures. It would be impossible for very many people, I think, to know very much about reactors in detail. But is there any hope—and perhaps the authors might comment on this—of simplified equations that would enable us to design reactors? Having produced a prototype, could we go on from one to another, with confidence that succeeding models would have predictable behaviour to limits which might be laid down by some body? The difficulty, as I see it after reading these papers, is how to marry the behaviour of the complex machine with the experiments we carry out on it. How can we develop the theory, whether by machine, by digital computer or by simplified methods, to enable us to continue building and improving the reactors which are an essential part of the country's economy?

Mr. A. Gilmour: In the concluding sentence in his paper, Dr. Howlett says that the needs of the nuclear-energy world will have provided an important part of the stimulus for the development of the digital computer. This is true, not only of the machines, but also of their use. For example, at a computing centre in which a digital computer has been used, during 1956 and 1957 some 27% of the time was spent on atomic-power problems, about 33% on aircraft-engineering problems and about 40% on other problems. When it is realized that the latter figure covers many branches of mechanical and electrical engineering, as well as commercial problems, it is seen that nuclear engineering is one of the greatest users of computers at this centre. However, the older branches of engineering have not been slow to follow the lead of the newer branches. One of the reasons for this is that many of the numerical techniques developed in nuclear engineering can be applied in other problems. For example, the solution of the diffusion equations described by Dr. Howlett has many similarities to the problem of the calculation of eddy-currents in busbars, and Monte Carlo techniques developed on computers for neutron transport problems can be used for various traffic-flow problems.

On the size of computers for nuclear-reactor design studies a very convincing argument has been made in favour of the biggest and fastest, but in using a very fast computer there is the danger that programmes may not be prepared as carefully as for a smaller one, so that, in practice, the difference in speed

between the two is not so large as one might expect. Wilkinson[*] states that optimum coded programmes are nearly always twice as fast as the same programmes on orthodox machines with the same length of major cycle, and in a large range of problems the gain is as much as a factor of ten. In other words, it is still necessary with a fast machine to pay careful attention to the numerical method used, the organization of the programme and the use of appropriate programming techniques if the necessary improvement in operating time is to be realized in comparison with a well-designed optimum coded programme on a smaller machine.

Professor A. Porter: The question of the respective roles of analogue and digital computers in the investigation of nuclear-reactor design is important. It would appear that the main distinction is that the digital computer can be applied to all cases where the problem can be stated specifically and where the parameters of the system are known. Many problems in nuclear physics fall into this category, and frequently comparatively high accuracy is required and also the capability of handling large amounts of information; the digital computer is particularly valuable in studying such problems. On the other hand, when the problem cannot be too well defined—and this is frequently the case when disturbances of an arbitrary nature arise—it appears that the analogue machine is particularly useful.

Dr. Howlett's paper considers, amongst other problems, the study of the neutron population in a reactor, and attention is thereby focused on the problem of neutron diffusion. The attention of the author is drawn to the work of Argyris in formulating aircraft-structure stress calculations as matrices and in the use of a high-speed computer. For example, in the determination of eigenvalues the computer itself forms part of an iterative process and the output of the machine provides the input for the next stage of iteration automatically. The Argyris method is well established as one of the most powerful that has yet been applied to the stressing of complex aircraft structures. Perhaps the method may be adapted to the study of neutron diffusion in non-homogeneous reactors.

I agree strongly with Dr. Howlett that from a national point of view it is important to develop the ultra-high-speed computer. Although several smaller machines with moderate capacity may appear at first sight to be adequate in the study of such complex problems as the 3-dimensional reactor, the ultra-high-speed computer is a necessity.

Mr. R. W. Sutton: The power output of a reactor is theoretically independent of its size—subject, of course, to the system being critical—and depends essentially on the rate at which heat can be transferred from the core. However, to get better heat transfer one requires a large heat-transfer surface area and a larger coolant-duct volume. This could lead to a reduction in the infinite multiplication factor, and it is therefore apparent that there may well be a conflict between satisfying the heat-transfer requirements and the nuclear design requirements. Dr. Howlett discusses only problems of nuclear design. Has he done any work on the optimization of plant design, taking into account both heat-transfer and nuclear requirements, and has he any comment on the value of such studies?

In tackling a problem of the kind in which Messrs. Bowen and Masters have been interested, one must first set up a mathematical model to represent the system to be examined. This is generally a set of ordinary differential and/or partial differential equations, and one has to obtain some solution of these. Errors between any computation carried out and the actual behaviour measured on the system can arise (a) from the inability of the model adequately to represent the system, (b) because one does not

know the data accurately, and (c) because the computing inaccurate. The obvious course is to try to eliminate all the errors in turn, so that, whatever the problem, the computation are as accurate as possible.

It is unsound to justify the use of an approximate solution on the ground that one does not know the data accurately cannot compute accurately. Inaccuracy in data can be covered by studying a range of values of a parameter, say, $\pm 10\%$ if the parameter is known only to that accuracy, to see whether th has a significant effect. One thus sees the significant parameter in the system, and can reduce the effort involved in determining parameters accurately.

As to the model used to represent the system, obviously th better the model the more nearly will the answers agree wit experiment. However, in order to make preliminary investiga tions on a system, one frequently wishes to investigate a wid range of parameters. From this aspect, methods such as thos suggested in the paper are extremely useful. However, there i always a danger that this type of approximation may be use rather indiscriminately. One does not always think carefull before using it in a particular case, to see whether it reall applies. If necessary, periodical checks should be made with more accurate computer. Moreover, when one has used thi type of approximation to find the range in which one is reall interested in a system, one should again go to a more accurat form of computation, including as many factors as one can, i order to make quite sure that nothing missed out in the approxi mations has any significant effect.

In producing Figs. 2 and 12, Messrs. Bowen and Master have apparently used eqn. (10) to represent the power, and the get a sudden jump in power which shows up clearly on both Figures at the initial condition. Had they used eqn. (11), woul they not have avoided this and got the power rising more in th manner of the analogue results?

Were Figs. 6(b) and 6(c) obtained on the analogue compute or by the approximate method? If the former, have the author used their approximate method on this type of problem, where they have rates of change of reactivity?

Mr. D. V. Wordsworth: Dr. Howlett deals very adequately with the steady-state nuclear problems which enter into th design of the reactor, and demonstrates clearly the high degre of complexity which arises in these calculations. There are two aspects of reactor design, however, which are not dealt with ir such detail, namely the thermodynamic design and the kineti behaviour of the system.

At the present stage of reactor design, considerable economi gains are possible by optimization of the relationship betweer heat output from the system and overall plant efficiency. Ar analysis of the possible steam cycles—examining the single- anc multi-stage pressure cycles and the various degrees of reheat—i an enormous task to attempt by desk calculation, and if this i included with an economic survey the task becomes formidable No doubt considerable improvements are possible here, and the economic gains for a given plant might well cover the cost o the computer required to achieve them.

It is realized now that consideration of the kinetic behaviou of the reactor must be taken into account in the early stages o the design study. The results obtained from such an analysi might be predominant in determining certain design features and kinetic studies should thus be of a very high standard Does Dr. Howlett believe that the problem could be dealt with adequately on a digital machine, bearing in mind the possibility of having to produce a multi-group, multi-region time-dependent calculation?

The paper by Messrs. Bowen and Masters illustrates some very interesting techniques and shows how the time behaviour of the

* WILKINSON, J. H.: 'An Assessment of the System of Optimum Coding used on the Pilot Automatic Computing Engine at the N.P.L.', *Philosophical Transactions of the Royal Society*, A, 1955, p. 253.

der Hall type of reactor can be analysed without the use of computer. It is necessary to stress, however, that the methods not general ones and should be applied with care to other tems. Attention should be drawn to certain assumptions de in the paper. First, the temperature coefficients are umed to be independent of temperature. Secondly, the mean utron lifetime is assumed to be independent of the reactivity the system. Thirdly, certain material properties, such as cific heat and heat-transfer coefficients, are assumed to be ependent of temperature. In some reactor designs, particu- ly in small systems operating at high temperatures, each of se will be highly variable in a transient. Temperature coeffi- nts might vary by a factor of five and heat-transfer coefficients a factor of two in a single excursion.

Mr. C. P. Gratton: Dr. Howlett deals admirably with the ny difficulties arising in the design of a nuclear power reactor h the increasing use of the high-speed general-purposes digital mputer. It is unfortunate, however, that in his choice of ample he has done little to dismiss the attitude so frequently opted. But this is a simple straightforward calculation; y put it on a machine?

Although more ambitious techniques may be used in reactor sign, the real need is for fast general approximate methods rforming general survey work. By this I mean fast both in achine sense and also for the designer who is controlling investigation. My experience in this field indicates that for ry hour on the machine approximately 100 are required to rmulate the problem, prepare the input data, and finally assess results. This ratio was obtained from the use of a wide nge of programmes covering kinetics, long-term changes with radiation, criticality and the evaluation of such things as mean utron lifetime and resonance neutron absorption integrals.

I feel that, although machines will be speeded up in the future, ey will not be quite so revolutionary to the designer as the recasts quoted in the paper suggests. However, the approxi- te methods I have mentioned would be poor indeed if more vanced techniques were not forthcoming to be used as a rdstick against which to check preliminary results and investi- te various approximations. It is in this respect that Dr. owlett's examples are essential, but I doubt whether the man- wer available could cope economically with general survey rk using these methods.

Mr. T. J. O'Neill: In the paper by Messrs. Bowen and Masters, me of the assumptions used in deriving the equations are ade to simplify the problem to a point where graphical calcula- n is possible. One major assumption made is that thermal diation from the fuel element to the moderator may be neg- cted. In some transients the gas flow falls to a low level, us reducing the heat transfer by convection, whereas the fuel- ement temperature increases, so increasing the radiative heat ansfer. It has been found in some cases that the heat transfer radiation exceeds that by convection by a factor of 5, and e course of the transient is markedly different when the radia- on term is included.

In deriving equations of this type it is necessary to assume at the shape of the gas-temperature curve along the channel mains constant with time; but in determining the mean tem- ratures it is not essential to assume that the shape of the fuel- ment-temperature curve is also constant. For the maximum peratures it is preferable to establish separate equations, ich require only the gas temperature to have an assumed m. These equations can then also include the appropriate iation terms, which are of different magnitude from those for mean equations.

he authors' equations neglect moderator-temperature changes the assumption that the time-constant is of the order of 30 min.

However, owing to the low thermal conductivity of irradiated graphite, there is a build-up of temperature near the edge of the moderator, with a much shorter time-constant. In addition, the reactivity temperature coefficient of the moderator is not uniform across the graphite, but is greatest nearer to the edge of the channel. The net effect is that a significant part of the reactivity produced by the moderator changes with a time- constant of approximately 2 min.

In applying the equations for one channel to the whole core, would one not expect to use a statistically weighted average effective channel for mean temperatures, rather than to select an actual channel at some radial point?

Mr. A. L. Gray: Messrs. Abson, Salmon and Pyrah mention that fission counters have applications in small reactors where low fluxes require the counter to be placed very close to the core. This also is the case in some of the larger graphite-moderated reactors, where the instrument engineer is required to use a location for his detectors which is some considerable distance from the core, and where the fluxes are low. The very wide range for which instrumentation is required means that low-level instrumentation must cope with neutron fluxes of the order of a few neutrons per square centimetre per second. At these levels good counting statistics are difficult to obtain, and there is a very strong case for operating counters inside the pressure vessel, if at all possible. The fission counter lends itself to applications of this type. The use of ceramic seals suggests that one can operate these counters at high temperatures, and I understand that a number of experiments have confirmed this.

A detector for use inside the core must be treated as a per- manent part of the installation and not require replacement at too frequent intervals. I should like the comments of the authors on life from two aspects. First there is the question of burn-up: here a simple calculation suggests that it might be possible to achieve lives of the order of 10 000 hours, which is quite reasonable, before burn-up becomes too serious a problem. Secondly, what is the upper limit of flux at which the counter may remain when the reactor is operating at full power, i.e. what limit is set by the effects of pile-up from residual β- and γ-activities in the structural materials of the counter?

Mr. D. J. E. Evans: I should like to ask Messrs. Bowen and Masters whether their study has indicated the degree of accuracy of speed maintenance needed on the blower drives.

Would Messrs. Abson, Salmon and Pyrah indicate the accuracy of production and the extent to which shelf life has a bearing on the active life of proportional counters? Do they consider it possible to develop fission counters, for flux plotting purposes in power reactors on load, having a better performance than the present-day wire-and-tape methods?

Mr. H. McGregor Ross (*communicated*): Dr. Howlett fails to show that there is a great range of applications concerned with the physics of the nuclear reactor which can readily be tackled with medium-sized digital computers. Examples of such work include computation of the main physical properties of the lattice, assessment of long-term reactivity, flux-distribution calculations and control of the flux level within the reactor. In addition to this work associated with the physics of the reactor, there is a wide range of applications in nuclear engineering. Much of this can benefit from programmes already written for other branches of engineering, e.g.

Heat-cycle optimization, including the prime mover.
Heat-exchanger calculations.
Stressing of the spherical pressure vessels, and their fittings.
Stressing of frameworks, especially to minimize deflections at prescribed points.
Stressing of pipe systems and bellows.
Centrifugal stressing in turbines and fans.
Critical speeds for the rotors of turbine-alternator systems.

For applications of this nature, experience shows that it is a great advantage for the design engineer to be able to prepare his own work for the computer. It is therefore of primary importance that the computer should be simple to programme, that the basic programming techniques should be fully worked out to give the maximum versatility and flexibility, and that there should be an extensive and easily used library of standard routines.

Another new advance of importance for this work is th Autocode schemes. These make a computer appear muc simplified, so that it is possible for engineers to learn in a da or two how to use it. Programmes are prepared in a sel evident notation. Practical Autocode schemes now exist fc the Deuce, Mark 1* and Pegasus computers, while that fc Mercury is distinguished by being very versatile and by pro serving the exceptional computing speed of the machine.

THE AUTHORS' REPLIES TO THE ABOVE DISCUSSION

Dr. J. Howlett (*in reply*): In reply to Mr. Farmer I would say that from the users' point of view digital computers are becoming simpler, and are providing facilities which take a lot of work off the programmer's shoulders; I agree that in the engineering sense they are becoming very elaborate. So far as understanding reactors is concerned, the impression I get from talking to my colleagues is that we have still a lot to learn: quite elaborate mathematical studies are needed to understand what goes on in a reactor, and any eventual simplification will have to be based on results of such studies.

I agree with Mr. Gilmour that all programmers must have a sound knowledge of numerical techniques and must bring a critical mind to bear on the choice of method, programming technique and general organization of the problem. While also agreeing that one must not too easily assume that only the largest and fastest machines are of real use, I feel that programming—which is really only a means to an end—should be as easy as possible, and the more elaborate machines contribute a lot to this aim.

I am grateful to Professor Porter for telling me of the work of Argyris on matrix processes. To Mr. Sutton I confess that I have not yet had to do any work on the optimization of a complete plant, but that I expect to be involved in calculations of this kind shortly. It seems very likely that, as the importance of economic questions in this field increases, such studies will become an essential part of any design project.

The time-dependent calculation suggested by Mr. Wordsworth can certainly be done on a digital machine, but it is a large undertaking if a 2-dimensional treatment is desired; I have, in fact, done something on these lines, but only a few steps in time were needed. I think that there is here no alternative to the digital machine: an analogue computer would be quicker, but one elaborate enough to deal with a complex problem would be extremely expensive.

I cannot agree with Mr. Gratton that the examples I give are 'simple straightforward calculations': even the simplest of the three (the spherical-harmonics calculation) is a considerable task for a hand computer, and the whole point of using a machine is that any of these calculations would take far too long if done by hand. Commenting on his later statement I would say that a most important aim is to make the machine do as much as possible of the assessing of the results of its computations—probably a more difficult task than writing the main programme.

Mr. Ross's list of desiderata for a machine seem to me wholly admirable. I am aware of the Autocode programming schemes and have made a great deal of use of this kind of technique over the past few years. I had, in fact, considered including something on programming, and in particular on this kind of project, in my paper, but it seemed to me that I had already written quite enough without this.

Finally, there is Dr. Taylor's question on faster machines: I consider it essential that a very advanced machine should be developed in this country. At the same time I think we must never cease to study the mathematical techniques on which large-scale calculations are based, or otherwise we shall not mak the best use of our machines, however fast they become.

Messrs. J. H. Bowen and **E. F. O. Masters** (*in reply*): Regardin Dr. Taylor's comments, although the methods employed in th paper may be applied to the analysis of such situations as th Chalk River accident and the American experiments with Borax Godiva, etc., making due allowance for the elasticity of th moderator in the case of water-moderated systems, to the bes of our knowledge there are no other directly applicable experi ments which have been made on graphite-moderated systems.

In reply to Mr. Sutton, we agree with his contention that logically, the use of an approximate solution is not justified b ignorance of the data involved. Figs. 2 and 12 are both draw for the condition of a sudden application or step of k_{ex}, wherea eqn. (11) is recommended only where reactivity is increased at constant rate.

We agree with Dr. Wordsworth that it is necessary to tes assumptions such as the variation of the parameters wit temperature. In the Calder-type system, all the assumption he has noted do apply with good accuracy. We are a littl surprised to hear of his experience of a five-fold variation of tem perature coefficient in a given transient. The basic phenomen summed up in the parameter 'temperature coefficient' are usuall proportional to the absolute temperature or fractional powe thereof, and such a variation as he reports would be expected onl over a very wide range of temperature. An artificial variatio is often noted, which is really a reflection on the suitability of th temperature coefficient chosen, due to, for example, the variatio of flux or of temperature distribution during the course of th investigation. Such factors are discussed in the paper.

In reply to Mr. O'Neill, the equations given do not allow fo radiant heat transfer from the fuel elements, although an approxi mate allowance may easily be made.

The diffusion of heat into the moderator can be considered i the simplified example shown in Fig. A, where the rise o temperature is plotted within a graphite slab, in contact on on face with a gas rising in temperature at a constant rate. The values of the constants assumed for the graphite take accoun of the irradiation effects mentioned by Mr. O'Neill. The slab surface temperature T_S follows the gas temperature T_A more closely than does the central temperature, and the effective temperature T_E which has been weighted for the greater effective ness of the surface moderator responds more quickly than that calculated for lumped parameters, T_L. If the graph were con tinued for longer times, the plots of T_A, T_E, and T_L would eventually become parallel. T_L would lag behind T_A by 10 min (the time-constant in this example), and T_E by about 6 min. However, as will be realized from the paper, practical tempera ture transients do not usually last longer than a few minutes; and although strictly it is correct to speak of a time-constant only if distribution effects are neglected, the simplification in the paper appears justified in that the thermal time-delay in the moderator is several orders of magnitude greater than that for the fuel.
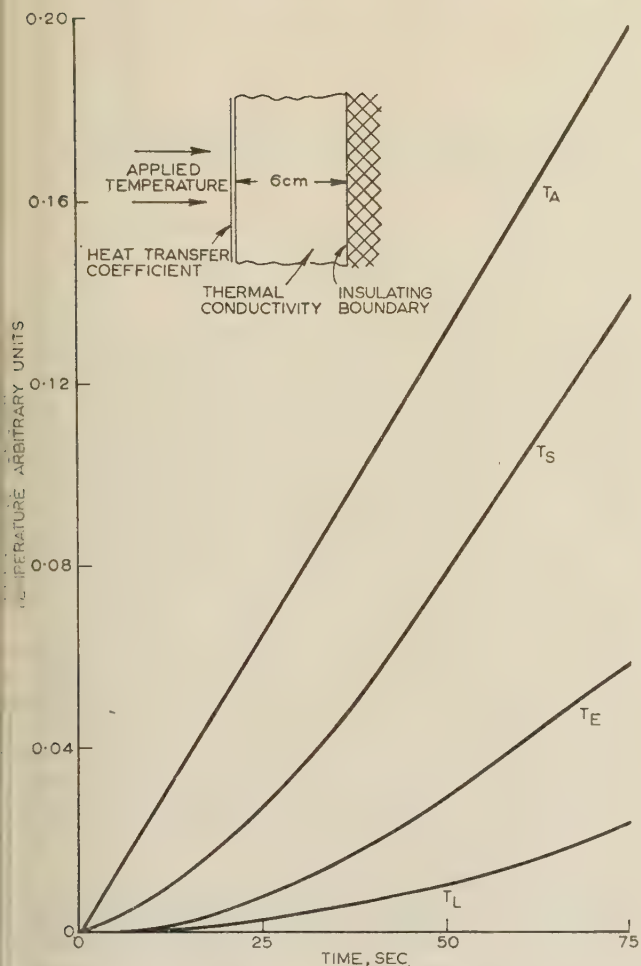
Fig. A.—Diffusion of heat into the moderator.

Heat transfer coefficient, $h = 4 \cdot 72 \times 10^{-3}$ C.G.S. units.
Thermal conductivity, $k = 0 \cdot 066$ C.G.S. units.
Density, $\rho = 1 \cdot 7$.
Specific heat, $C = 0 \cdot 3$.

Regarding allowance for radial variations, the procedure recommended is, in fact, to use a statistically weighted average effective channel for mean temperatures, exactly as Mr. O'Neill suggests. By the procedure given in the paper, the actual channel whose behaviour corresponds to this average may be identified. The behaviour in any other channel may be calculated from this.

In reply to Mr. Evans, the methods used in the paper could be adapted to study the problem of the speed stability required on blower drives; but for this purpose, where interest is fixed on small amplitude changes, more conventional linear perturbation analysis may be employed.

Messrs. W. Abson, P. G. Salmon and **S. Pyrah** (*in reply*): In reply to Dr. Taylor, there is no evidence that fluorine is pro-

duced when boron trifluoride counters are heated. Fig. 8 refers to the deterioration of counting characteristics after a large number of counts, and it is postulated that this may be due to the production of free fluorine following dissociation of boron trifluoride in the counter discharge process. Subsequent heating of the counters gives rise to complete recovery, and this is consistent with absorption of the fluorine on the copper walls of the counter.

The effect of induced activities can be considerably reduced by using aluminium in place of copper for the main body of the counter, since the activation cross-section is reduced by an order of magnitude and the half-life of the induced activity of aluminium 28 is $2 \cdot 3$ min, compared with $12 \cdot 8$ h for copper 64. Boron trifluoride proportional counters with aluminium bodies have been made at the Atomic Energy Research Establishment, but the reject rate has been high and the results on shelf-life have been variable. This is almost certainly due to the increased difficulties of avoiding contamination of the boron trifluoride with impurities from the aluminium walls and in obtaining a really effective vacuum seal between the aluminium and the nickel-iron of the glass-metal lead-through seal. An alternative approach to this problem is to use a boron-coated cathode, with argon and carbon dioxide as the gas filling. Some work is proceeding at Harwell on a counter of this type, and initial results on counting life at high counting rates are showing promising results, but data on shelf life are not yet available.

Fission counters of the type shown in Fig. 11 have been operated satisfactorily at 350° C, above which some difficulties have been encountered with spurious counts due to small breakdown pulses in the metal-ceramic seal. This may be due either to stresses in the ceramic or to changes in its insulating properties after extended use at high temperatures. The assembly details are being modified to reduce stresses due to relative expansions, and higher-purity alumina ceramics are being investigated.

In reply to Mr. Gray, the upper limit of neutron flux at which a fission counter may be exposed without limiting its subsequent use due to induced activities will depend on the size of the counter and the materials of construction. An estimate of the induced-activity effects in the B165 counter described (stainless-steel cathode of area 165 cm²) shows that the limit is greater than 10^{11} n/cm²/sec; pulse build-up effects due to induced beta activity at this flux level would be comparable with that due to 10^5 r/h of γ-radiation.

In reply to Mr. Evans we consider it unlikely that pulse-type fission counters could be made for flux plotting in thermal reactor cores at flux levels above 10^{11} n/cm²/sec. Above this level the pulse build-up effect of the accompanying γ-radiations would be prohibitive. A more promising possibility for flux measurements in the core of a power reactor on load, with flux levels of the order of 10^{13} n/cm²/sec, is the use of a d.c. ionization chamber with an electrode coating of boron or of a fissile material such as uranium 235. The main problem—which has not been completely resolved—is the design of a flexible cable having adequate insulation resistance at high temperatures and high radiation levels.

THE SCREEN EFFICIENCY OF SEALED-OFF HIGH-SPEED-OSCILLOGRAPH CATHODE-RAY TUBES

By R. FEINBERG, Dr.Ing., M.Sc., Associate Member.

SUMMARY

The factors which affect the screen efficiency of a high-speed-oscillograph cathode-ray tube are summarized, and the nature of the shape of the luminance pulse produced by a short-duration square-wave screen excitation is explained. It is then shown how shortening the duration of screen excitation reduces the screen efficiency of the tube as a consequence of energy lost by non-radiative dissipation.

(1) INTRODUCTION

The screen efficiency of a cathode-ray tube designates the ratio between the luminescent radiation emitted from the tube and the electrical power carried in the electron beam to excite the screen; it depends on a number of factors inherent in the design, construction and operation of the tube, namely

(*a*) The intrinsic luminescent efficiency of the phosphor employed to make the screen.
(*b*) The thickness and physical structure of the screen, which consists either of a texture of crystalline phosphor-powder particles or of a semi-transparent phosphor film.
(*c*) The method of screen laying.
(*d*) The absence or presence of a metal film at the back of the screen.
(*e*) The condition of screen excitation, i.e. electron velocity and current density of the electron beam at the screen, duration of the screen excitation and any previous excitation.
(*f*) The transmission properties of the cathode-ray-tube face.
(*g*) The spectral distribution of luminescent radiation.
(*h*) The properties of the optical system and of the photographic material if the cathode-ray oscillograph operates with photographic recording of the trace.

The efficiency of screens with phosphor crystals was investigated by Hopkinson[1] in respect of the influence of phosphor particle size, screen thickness, method of screen laying, condition of screen excitation, and various types of phosphor. Bril and Klasens[2] determined the intrinsic efficiency of a number of crystalline phosphors and measured also the increase of screen efficiency caused by an aluminium film at the back of the screen.

It is of interest to examine the particular conditions affecting the screen efficiency of a cathode-ray tube used for high-speed oscillography.

(2) MECHANISM OF LUMINESCENT RADIATION IN HIGH-SPEED OSCILLOGRAPH PHOSPHORS

The phenomenon of luminescence in a phosphor is now generally explained with the concept that a phosphor is based on a pure insulating crystal which is made luminescent by the addition of a small proportion of so-called 'impurity atoms' which are supposed to occupy interstitial or substitutional positions in the matrix lattice. Impurity atoms or other irregularities in the crystal lattice give rise to localized energy states of electrons in the crystal which are responsible for the mechanism of luminescence when the phosphor is excited, either by irradiation of electromagnetic energy or by electron bombardment as

in a cathode-ray tube. Two types of localized electron energy states are believed to exist in a phosphor crystal, namely the so-called 'luminescence centres' at which excited electrons—the result of energy absorption during phosphor excitation—return to the ground state, thus causing the emission of luminescent radiation, and the so-called 'electron traps', which are capable of capturing electrons excited to the energy level of the conduction state. A trap holds a captured electron until its release as a consequence of thermal agitation, and electrons released from electron traps may either be retrapped or may move to a luminescence centre to return there to the ground state with a resultant emission of luminescent radiation.

The mean time interval between electron excitation and return to the ground state, i.e. the mean life-time of an excited electron in the phosphor, determines the persistence-decay characteristic of the phosphor. This is exponential and has a short time-constant in those phosphors where the excited electrons spend their life-time in luminescence centres without being trapped, but it is complex and the decay time may be long if the phosphor has such a structure that excited electrons are trapped before returning to the ground state.

A persistence-decay characteristic which has been obtained experimentally is shown in Fig. 1, the experimental technique being described in Section 7. The phosphor investigated is a blue-radiating silver-activated zinc-sulphide of a type used for high-speed-oscillograph cathode-ray tubes. Fig. 1 is plotted on a semi-logarithmic co-ordinate system to facilitate an analysis of its character: the linear section demonstrates an exponential decay with a short time-constant, which indicates that excited electrons spend their life-time in luminescence centres without passing through an electron trap, and the non-linear section, for the first 3–4 microsec of persistence, shows that during this time a small fraction of luminance is produced by electrons which were trapped after excitation. The shaded area in Fig. 1 is a measure of the magnitude of the total effect of electron trapping.

(3) PHOSPHOR PERFORMANCE IN SHORT-PULSE OPERATION

Fig. 2 shows beam-current square-wave pulses of various durations but uniform magnitude and the corresponding screen-luminance pulses obtained with the cathode-ray tube (and hence the same phosphor) used to obtain the persistence characteristic shown in Fig. 1. The accelerator voltage of the tube, which had an aluminized screen, was 15 kV and the beam current-density was about 5 mA/cm². The luminance curves (*a*), (*b*) and (*c*) in Fig. 2 correspond to beam-pulse durations of 30, 10 and 2 microsec, respectively; they show that the screen luminance rises exponentially, reaching a steady-state value if the beam-pulse duration is sufficiently long, but otherwise being chopped off during the rise when the beam-pulse duration is short. The decay sections also obey the same law as the persistence characteristic, which is thus independent of the duration of phosphor excitation. The experimental technique employed to obtain the curves of Fig. 2 is described in Section 7.
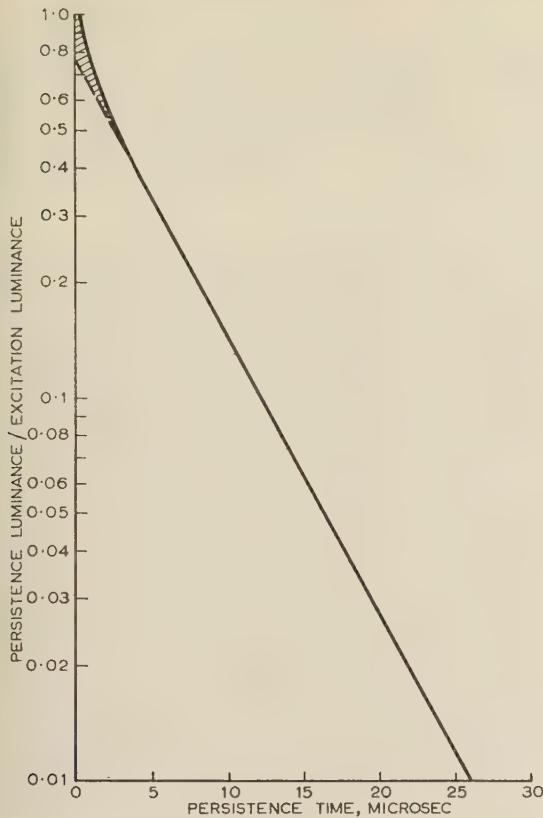
Fig. 1.—Experimentally obtained persistence-decay characteristic of a blue-radiating silver-activated zinc-sulphide phosphor used for high-speed-oscillograph cathode-ray tubes.
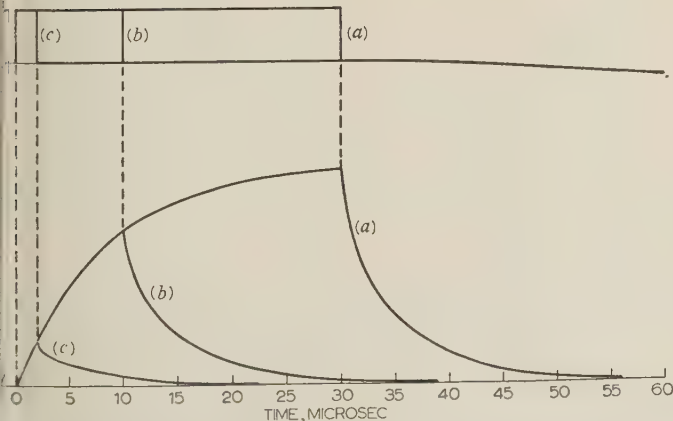
Fig. 2.—Experimentally obtained luminance pulses for various durations of phosphor excitation with the same phosphor as for Fig. 1.

    (a) 30 microsec pulse.
    (b) 10 microsec pulse.
    (c) 2 microsec pulse.

The rise and decay of luminance are the result of energy storage during the luminance rise time and release during the decay time. Energy is stored in the form of excited electrons which, because of their short life-times, are continuously replenished. It is of interest to consider the energy relationship
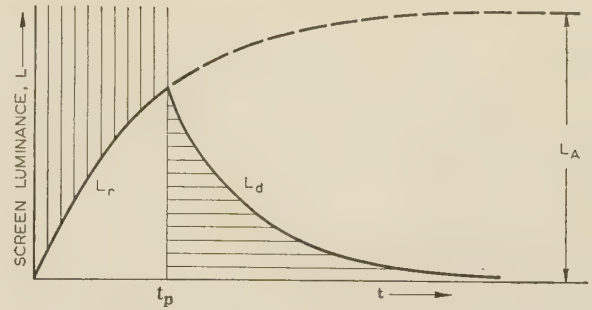
Fig. 3.—Luminance pulse with idealized persistence decay.

between storage and release. Fig. 3 shows a slightly idealized luminance pulse based on the luminance curves in Fig. 2. The luminance rise, L_r, is exponential, in line with Fig. 2, and the luminance decay, L_d, is also exponential, but is an approximation compared with Fig. 2. Let L_A be the steady-state luminance, t_p the duration of the beam-current pulse and hence of phosphor excitation, and τ_r and τ_d the time-constants of luminance rise and decay.

Then
$$L_r = L_A(1 - \varepsilon^{-t/\tau_r}) \quad . \quad . \quad . \quad . \quad . \quad (1)$$

and
$$L_d = L_A(1 - \varepsilon^{-t_p/\tau_p})\varepsilon^{-(t-t_p)/\tau_d} \quad . \quad . \quad . \quad (2)$$

The energy, W_1, stored in excited electrons may be considered to be proportional to the vertically shaded area in Fig. 3. Thus, from eqn. (1),

$$W_1 = \int_0^{t_p} (L_A - L_r)dt$$

$$= L_A\tau_r(1 - \varepsilon^{-t_p/\tau_r}) \quad . \quad . \quad . \quad . \quad (3)$$

The energy, W_2, released during decay is proportional to the horizontally shaded area in Fig. 3. Therefore, from eqn. (2),

$$W_2 = \int_{t_p}^{\infty} L_d dt$$

$$= L_A\tau_d(1 - \varepsilon^{-t_p/\tau_r}) \quad . \quad . \quad . \quad (4)$$

Thus, from eqns. (3) and (4),

$$W_2/W_1 = \tau_d/\tau_r \quad . \quad . \quad . \quad . \quad . \quad (5)$$

In practice we have $\tau_d < \tau_r$; therefore,

$$W_2 < W_1 \quad . \quad . \quad . \quad . \quad . \quad . \quad (6)$$

The difference between W_1 and W_2 is energy lost in the form of non-radiative dissipation.

In Fig. 2 $\tau_r = 8\cdot4$ microsec and, for the exponential section of the luminance decay, $\tau_d = 6\cdot0$ microsec; therefore $W_2/W_1 = 0\cdot7$. Considering the approximation postulated to derive eqn. (5) we find that somewhat more than 30% of the energy stored during luminance rise is lost in non-radiative dissipation.

(4) SCREEN EFFICIENCY IN HIGH-SPEED-OSCILLOGRAPH OPERATION

The spot velocity, v, on the screen of a cathode-ray tube and the spot diameter, d, determine the duration, t_p, of phosphor excitation during a single sweep of the electron beam across the screen, i.e. $t_p = d/v$. The minimum spot velocity under given conditions of tube operation occurs when no signal deflection is applied. Consider, for example, a tube operating with a time-base of 100 mm length and a linear sweep time of 1 microsec, and with a spot diameter of $0\cdot3$ mm. The minimum spot velocity then becomes 100 mm/microsec and the maximum phosphor

excitation time is $0 \cdot 003$ microsec. A phosphor with luminance rise and decay characteristics as shown in Fig. 2 produces at such short phosphor excitation times a luminance pulse with a shape probably somewhat similar to curve (c) in Fig. 2.

The luminance curves in Fig. 2 demonstrate that reducing the duration of phosphor excitation results in an increase of the ratio of luminescent radiation during the decay time to that during the rise time. This increase in the ratio reduces the screen efficiency, because the energy released as luminescent radiation during the decay time is smaller than that stored in excited electrons during the rise time [see eqn. (6)]. Let W_3 be the energy of luminescent radiation of a luminance pulse, and W_4 the respective energy supplied.

Then
$$W_3 = \int_0^{t_p} L_r dt + \int_{t_p}^{\infty} L_d dt$$

$$= L_A[t_p - (\tau_r - \tau_d)(1 - \varepsilon^{-t_p/\tau_r})] \quad . \quad . \quad (7)$$

and
$$W_4 = \int_0^{t_p} L_A dt$$

$$= L_A t_p \quad . \quad . \quad . \quad . \quad . \quad . \quad (8)$$

Hence
$$W_3/W_4 = 1 - \frac{(1 - \tau_d/\tau_r)(1 - \varepsilon^{-t_p/\tau_r})}{t_p/\tau_r} \quad . \quad . \quad (9)$$

For $t_p \ll \tau_r$, eqn. (9) becomes

$$W_3/W_4 = \tau_d/\tau_r \quad . \quad . \quad . \quad . \quad . \quad (10)$$

The ratio τ_d/τ_r is the measure of the loss of energy between storage and release.

In the example of the cathode-ray tube under consideration we have $\tau_d/\tau_r = 0 \cdot 7, \tau_r = 8 \cdot 4$ microsec and $t_p = 0 \cdot 003$ microsec. Thus, eqn. (10) applies, and we obtain $W_3/W_4 = 0 \cdot 7$. Since the calculation has been made by approximation, the result means that the short excitation because of the high spot velocity reduces the efficiency of the phosphor to below 70% of the value at steady-state excitation.

The shape of the luminance pulse may be very different from curve (c) in Fig. 2 when the spot velocity is very high, say 5×10^5 mm/microsec or more, as was obtained, for example, with a spot diameter of $0 \cdot 5$ mm in an oscillograph cathode-ray tube of a particular design[3] also containing a screen phosphor of the type used in the tube with which the curves in Figs. 1 and 2 were obtained. At $v = 5 \times 10^5$ mm/microsec and $d = 0 \cdot 5$ mm the duration of phosphor excitation during a single sweep of the beam is 1×10^{-6} microsec, which means that luminescent radiation probably starts virtually abruptly and when the excitation has already ceased. Without any further experimental evidence it is not possible to state how the efficiency of the phosphor is affected by the shape of the luminance pulse in the case of very high spot velocity.

In addition to the effects discussed, the screen efficiency of a high-speed oscillograph cathode-ray tube is further affected by the behaviour of the photographic material used to record the trace. The subject of photography at very short exposure times is complex and not yet clarified.[4,5] It is therefore not possible at present to complete the picture of the factors which determine the screen efficiency of a high-speed-oscillograph cathode-ray tube.

(5) CONCLUSIONS

The particular conditions which affect the screen efficiency of a cathode-ray tube employed for high-speed oscillography are the reduction of the efficiency of the phosphor as a consequence of a certain non-radiative dissipation of energy becoming significant at high spot velocities, and the behaviour at very short exposure times of the photographic material used to record the trace.

(6) REFERENCES

(1) HOPKINSON, R. G.: 'An Examination of Cathode-Ray-Tube Screen Characteristics', *Journal I.E.E.*, 1946, **93**, Part IIIA p. 50.

(2) BRIL, A., and KLASENS, H. A.: 'Intrinsic Efficiencies of Phosphors under Cathode-Ray Excitation', *Philips Research Reports*, 1952, **7**, p. 401.

(3) JACKSON, B., HARDY, D. R., and FEINBERG, R.: 'A Sealed-off Cathode-Ray Tube with a Very High Writing Speed', *Nature*, 1953, **172**, p. 1056.

(4) FRIESER, H.: 'Eigenschaften photographischer Schichten bei kurzen Belichtungszeiten', *Actes du Deuxième Congrès International de Photographie et Cinématographie Ultra-Rapides*, Paris, 1956, p. 184.

(5) CASTLE, J., WOODBURY, W., and SHELTON, W. A.: 'Reciprocity-law Failure at Very Short Exposure Times', *Proceedings of the Third International Congress on High-Speed Photography*, London, 1957, p. 219.

(6) DUDLEY, M. D., WHITNEY, I., and FEINBERG, R.: 'A High-Intensity Light Source for Stroboscopic Operation at Frequencies up to 250 kc/s', *Nature*, 1957, **179**, p. 1024.

(7) APPENDIX

The method applied in the experimental technique used to obtain the curves in Figs. 1 and 2 was based on a cathode-ray tube of a suitable design and with a screen made of the phosphor to be investigated. The screen was excited with a recurrent square-wave beam current pulse of known magnitude and duration, and an electron-multiplier photocell converted the luminance pulse from the screen into a corresponding current pulse which controlled a wide-band amplifier whose output voltage pulse was displayed on the screen of a high-speed cathode-ray oscillograph. The time-base of the oscillograph was triggered in synchronism and appropriate phase relationship with the beam current pulse of the cathode-ray tube under investigation. It was convenient for the operation of the system to use the mains supply as the source of recurrence frequency, giving sufficiently long intervals of about 20 millisec between successive pulses.

The cathode-ray tube employed for the investigation of the phosphor performance was of a type operating with heavy beam current pulses, the beam exciting the entire screen simultaneously and uniformly.[6] The magnitude of the beam current pulse was 100 mA, giving a current density of about 5 mA/cm² at the screen. For obtaining the curve shown in Fig. 1 the beam current-pulse duration was about 100 microsec, to ensure a steady-state condition of luminance at the end of each excitation pulse. For the curves in Fig. 2 the pulse durations were respectively 30, 10 and 2 microsec.

MAGNETIC TAPE FOR DATA RECORDING

By C. D. MEE, Ph.D., A.Inst.P.

SUMMARY

The occurrence of mistakes, i.e. drop-outs, in the recording and reproduction of pulse signals on magnetic tape is investigated. The mechanisms of recording and reproduction are discussed for return-to-zero and non-return-to-zero recording, and their susceptibility to drop-outs is assessed. Conditions for minimizing the occurrence of drop-outs using conventional recording methods are considered, and new recording systems are suggested and described which are insensitive to the majority of drop-outs.

An experimental test apparatus for detecting drop-outs and measuring their size is described which may be used to test tape under widely varying conditions of recording. This is used to study the effects of different plastic base materials on tape performance for data recording. The drop-out content of present-day tapes is measured, and an economical performance specification is suggested.

(1) INTRODUCTION

The major weakness in using magnetic tape as a storage medium in data-handling equipment is its susceptibility to errors due to momentary imperfections in the recording or reproducing processes. One purpose of the paper is to investigate how far these imperfections may be related to faults in the magnetic tape or to the system of recording and reproducing. For successful operation, tape faults should preferably not occur, although it is possible to avoid them in recording if their location is known. It is therefore necessary to test tape in a way which will locate possible faults for any system of data recording, and a suitable test apparatus is described.

Whatever method of recording or reproducing is used in an application of tape for data recording, the aim is to reproduce faithfully the number of pulses in the signal to be recorded and also to maintain their time sequence. In other words, the signal, consisting of a series of pulses spaced in time according to some predetermined code, must be reproduced without any loss or addition of pulses. By far the most common fault occurring in magnetic recording of pulses is one which momentarily separates the tape from the recording or the reproducing head, thus causing loss of amplitude of the reproducing signal. In data recording it is only necessary to know whether a pulse exists or not at any particular time. Hence a reduction in amplitude of the reproducing signal can be tolerated, provided that the pulses can be distinguished absolutely from the accompanying noise. When, for a given recording and reproducing system, the reproducing signal falls below this limit, a faulty reproduction of the original signal occurs, and this is called a 'drop-out'. Several data-recording systems are at present in vogue, and in order to guarantee freedom from drop-outs the testing method must include conditions of recording and reproducing similar to those to be used in practice, even though the latter may not always produce a minimum of drop-outs.

(2) DATA-RECORDING SYSTEMS

(2.1) Recorded Pulse Density

It is the aim of all data-recording systems to use the maximum pulse packing density on the tape compatible with absolute

reliability. This has led to the use of multi-track recordings, and a system in common use requires a tape width of 0·045 in for each recording, with 0·015 in between recordings. Other track dimensions are also used, and, since the effect of a tape imperfection on a recording depends on its lateral position with respect to the recording, it is important that the track dimensions used in the test apparatus should conform with those used in practice.

In each recorded track, the main limitation on the number of pulses per unit length of track is the increase of variations in the reproducing-signal amplitude with pulse packing density. However, with some systems, the inability to distinguish a pulse from its neighbour also limits the packing density. At present, between 100 and 200 pulses per inch of tape per track can be used satisfactorily.

(2.2) The Recording System

(2.2.1) Non-Return-to-Zero Recording.

The recording system having widest application is one which saturates the magnetic oxide in the longitudinal direction of the tape. The direction of magnetization changes through 180° in accordance with direction changes of the field in the gap of the recording head. The changes in head-field direction are usually rapid, and the spread of the transition region of magnetization on the tape is determined largely by the dimensions of the recording-head gap, the magnetic characteristics of the tape, and the separation between the head and the tape. Fig. 1 illustrates the tape and head dimensions relevant to the discussion of conventional longitudinal recording and reproducing systems. Up to 400 reversals of magnetization direction per inch of tape are used [Fig. 2(a)], and such a recording will thus consist of 400 magnets per inch, alternately magnetized to maximum remanent induction, separated by transition regions where the magnetization changes rapidly [Fig. 2(b)]. Since a surface magnetic field exists at the transition regions, their extent may be made visible with a field indicator such as colloidal magnetite. A recording similar to that shown in Fig. 2(b) is reproduced in Fig. 3, in which the black lines of magnetite deposits indicate that the transition regions are slightly narrower than the fully magnetized regions.

Two conflicting requirements emerge when considering the design of the recording head for the type of recording described (called non-return to zero). In order to reduce the separation loss when a pimple on the tape momentarily lifts the tape away from the gap, it is desirable that the rate of extinction of the recording field be as small as possible in the direction of separation. This is achieved with a long recording-head gap, gap length being measured in the direction of tape travel. On the other hand, in order to produce sharply defined pulses it is necessary to minimize the spread of the transition region on the tape between the two states of maximum remanent induction. This is achieved with maximum rate of extinction of the longitudinal recording field from the saturating value. The maximum rate of extinction occurs near the trailing edge of the gap, and in order to record at this point the field must be sufficient to saturate the tape at the trailing edge, otherwise recording will take place between the pole-pieces of the head, where the rate of extinction may be low. The rate of extinction will be related to

Dr. Mee was formerly with the M.S.S. Recording Co., Ltd., and is now with C.B.S. Laboratories, New York.
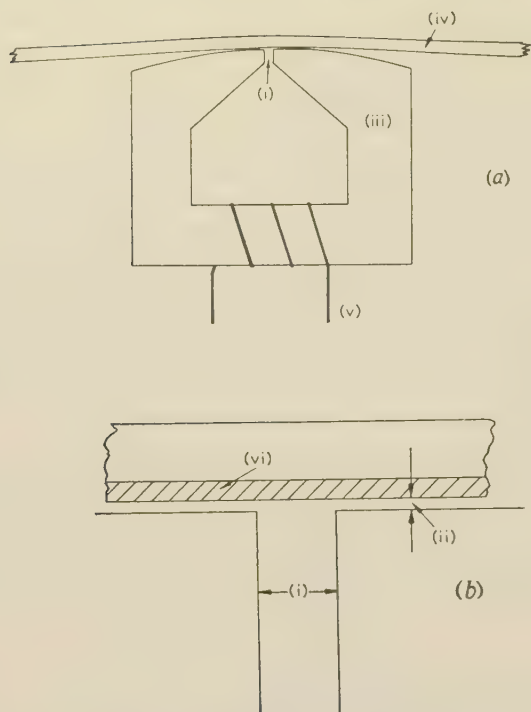
Fig. 1.—Definitions used in the paper.

(i) Head gap.   (ii) Head/tape separation.
(iii) Head core.   (iv) Tape.
(v) Head winding.   (vi) Magnetic layer of tape.

(a) Recording or reproducing head with tape.
(b) Enlargement showing head gap and tape.
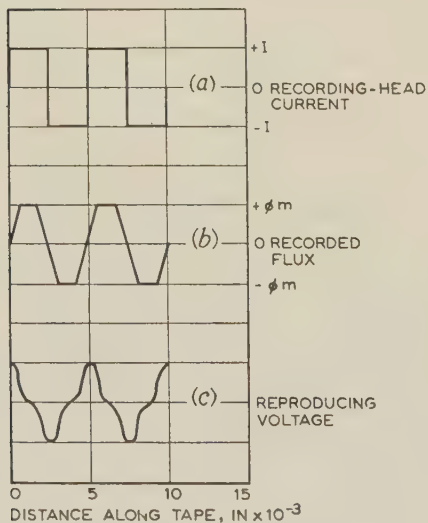
Fig. 2.—Waveforms for non-return-to-zero recording.

Fig. 3.—Non-return-to-zero recording.
Recording-head gap length = 0·001 in.
Magnification × 26.
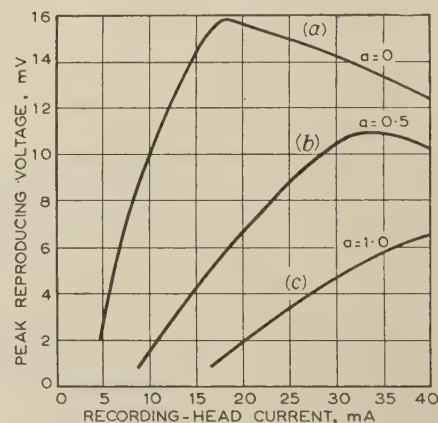
Fig. 4.—Recording losses due to separation (n.r.z.).
100 pulses per inch.
Tape length between reversals of magnetization = 0·005 in.
a = Head/tape separation, in × 10⁻³.
Gap length = 0·001 in.

the gap length and will be higher for small gaps. Hence, some compromise must be made between the desirable features of sharp pulses and minimum susceptibility to separation loss.

In order to determine the separation loss appropriate to a 0·001 in gap length, the relationship between the reproducing voltage and the recording current for a recording signal, similar in form to that shown in Fig. 2(a), was determined for different spacings between the recording head and the magnetic layer of the tape. The curves, appropriate to 200 reversals of magnetization per inch of tape, are shown in Fig. 4. The maximum

reproducing voltage occurs when the recording field at the trailing edge is sufficient to saturate the magnetic layer throughout its thickness (0·000 5 in) and when the rate of longitudinal extinction of the recording field below that required to saturate the tape is a maximum.

It is seen from Fig. 4 that a drop-out, which separates the head and tape by a distance equal to the oxide thickness, causes minimum loss of reproducing signal for recording currents in excess of twice that required to produce normal maximum output. For the latter recording current, the separation loss is about four times greater. Drop-outs of this magnitude are a maximum that can be tolerated due to separation loss on reproduction, and so a figure of double the recording current for normal maximum output may be taken as optimum for the conditions considered. For smaller gap lengths, the rate of extinction of the recording field with separation from the head is greater, and relatively larger recording currents would be required to produce a minimum separation loss.

(2.2.2) Return-to-Zero Recording.

As the name suggests, with return-to-zero recording the tape is unmagnetized in the regions between the recording pulses. The recording consists of pulses of either positive or negative magnetizing fields, which are short in duration compared with the
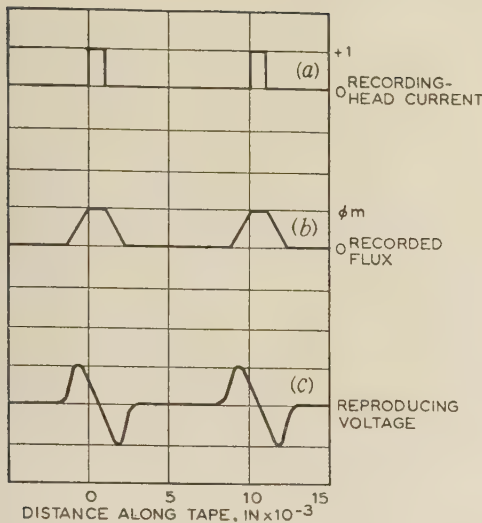
Fig. 5.—Waveforms for return-to-zero recording.

recording current above that required to produce maximum reproducing signal, i.e. saturation of the tape. The upper half of Fig. 6 corresponds to a recording current just sufficient to saturate the tape, and the lower half, to twice this current. An increase of recorded pulse length from $0 \cdot 002\,5$ in to $0 \cdot 003\,5$ in results from this increase of recording current. Since the recording pulse has a finite rise-time, an increase of recording current will result in tape saturation for a longer time and, in addition, the recording field will be lengthened beyond the limits of the recording-head gap. These effects account for the increase of recorded pulse length.

The spread of the recorded pulse sets a limit to the pulse packing density when an r.t.z. system of recording is used, and three states of magnetization have to be distinguished. In Fig. 7 the peak reproducing voltage is plotted as a function of

epetition period. A typical return-to-zero (r.t.z.) signal consisting of unidirectional pulses is shown in Fig. 5(a), where the recording-head magnetizing current, during the pulse, is sufficient to saturate the tape. The resulting colloidal magnetite patterns of this type of recording are shown in Fig. 6.
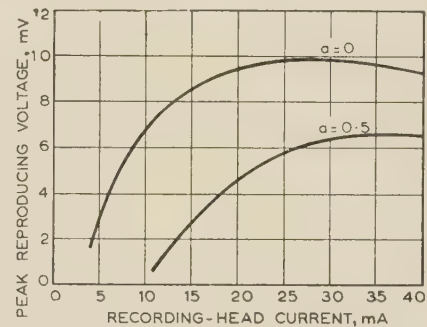
Fig. 7.—Recording losses due to separation (r.t.z.).

100 pulses per inch.
Pulse length on tape $= 0 \cdot 000\,5$ in.
a = Head/tape separation, in $\times 10^{-3}$.
Gap length $= 0 \cdot 001$ in.

recording current for zero and $0 \cdot 000\,5$ in head-to-tape separation. It is seen, by comparison with Fig. 4, that approximately the same peak recording current is required to minimize the loss due to $0 \cdot 000\,5$ in separation. Colloidal magnetite patterns show for this recording current that the length of the magnetized pulse for the recording conditions described is about $0 \cdot 003$ in. With such a recording, about 100 pulses per inch may be accommodated without interference between adjacent pulses. It would be expected that shorter recording-head gaps would produce shorter recorded pulses. However, the resulting restriction of longitudinal field is small when sufficient field is produced to saturate the tape. A shorter pulse is also produced when the length of tape passing the gap during the time of the pulse is negligible compared with the gap length. In this case the length of the recorded pulse for a $0 \cdot 001$ in recording-head gap is about $0 \cdot 002$ in.

Fig. 6.—Return-to-zero recordings.

Recording-head gap length $= 0 \cdot 001$ in.
Magnification $\times 26$.

The mechanism of recording with a signal of the type shown in Fig. 5(a) is different from that described in the previous Section. The length of the pulse of magnetization on the tape will depend on the length of the recording-head gap, and also on the length of tape which traverses the gap during the time that the pulse field is on. In this case, recording takes place across the whole length of the gap, rather than just at the trailing edge as described for non-return-to-zero (n.r.z.) recording, and hence a minimum length of the recorded pulse is set by the length of the recording-head gap. In the experiments to be described a gap length of 001 in and a pulse length along the tape of about $0 \cdot 001$ in are used. The pulse repetition length is $0 \cdot 01$ in, and the tape is therefore subjected to 100 pulses of saturating recording field per inch, where each pulse extends about $0 \cdot 002$ in along the tape.

It is observed from colloidal magnetite patterns (Fig. 6) that the length of the recorded pulse increases with an increase of

(2.3) The Reproducing System

(2.3.1) Conventional Contact Reproducing.

It is a general aim in data recording to attempt to obtain a minimum spread of the reproducing voltage pulses, so that the pulse packing on the tape is not limited by the difficulty of distinguishing one pulse from its neighbour. The reproducing system response will now be considered in the light of the desirable features of minimum sensitivity to drop-outs, along with the preservation of reproducing voltage pulses which are large in amplitude and short in duration.

The overall wavelength response of the reproducing chain will depend on the head-to-tape separation, the gap length and the frequency response of the reproducing system. It has been shown that the general design of the head itself also has a marked effect,[1] but this discussion will be confined to standard ring-type heads, since these at present have universal application, and the

main purpose of this investigation is to establish a method of testing data-recording tape using present techniques of recording and reproduction.

For an n.r.z. recording the reproducing voltage waveform is of the form shown in Fig. 2(c), and for an r.t.z. recording it is similar to that shown in Fig. 5(c). Increasing the short-wavelength response of the reproducing system improves the sharpness of the voltage pulses, and a limit to this is set only by the maximum rate of change of the magnetic flux in the reproducing-head core. In order to achieve sharpness of the reproducing pulses it is necessary to ensure that the reproducing head does not introduce high losses for the maximum tape surface flux-density gradient. The maximum will occur at the surface of the tape in the middle of the magnetization transition region. Losses will be small if the pole-tips of the head are separated by a distance (i.e. gap length) which is less than the total transition length and if they make good contact with the tape. When the tape is separated from the reproducing head, rapid initial decrease occurs in the surface flux intercepted by the heads, owing to loss of the short-wavelength components of the surface flux. The loss of reproducing voltage with separation, for the two systems of recording previously described, is shown in Fig. 8, curves (a) and (b), for a
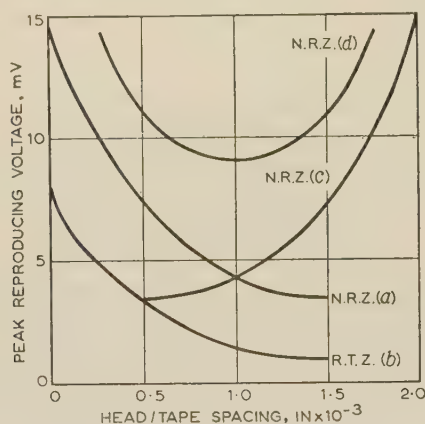
Fig. 8.—Separation losses on replay.
Recorded signal 100 pulses per inch.
Reproducing-head gap length = 0·000 5 in.

gap length of 0·000 5 in. In both cases the repetition length of the recorded signal is 0·01 in.

The use of a somewhat longer gap in the reproducing head does not materially alter the reproducing voltage, since the resulting loss of the short-wavelength component of the signal is partially compensated by the increased sensitivity of the head due to its greater gap reluctance. In addition, since the short-wavelength components are thus attenuated, the initial separation loss is less severe than in the previous case and drop-outs are less apparent.

In practice, some applications demand maximum sharpness of the reproducing pulse. Good contact between the tape and the head is essential in this case. A reproducing-head gap length not greater than half the transition length is used, and the bandwidth of the reproducing amplifier is wide enough to avoid attenuation of the high-frequency components of the reproducing signal. On the other hand, some applications can benefit by accepting a lower definition of reproducing voltage pulse, and in these cases a reduction in sensitivity to drop-outs can be achieved by deliberate separation of the tape from the reproducing-head pole-pieces. A second advantage is that a reduction of wear on the reproducing head is obtained. Non-contact systems will now be considered in more detail.

(2.3.2) Non-Contact Methods of Reproduction.

Observations of the causes of drop-outs with conventional recording and reproducing systems indicate that the majority are caused by momentary separation of the tape from the head. There are many different reasons for this separation loss, but in nearly all cases the layer of magnetic material is intact and capable of being magnetized in the normal manner. Hence, any recording or reproducing-head system which does not have the usual separation-loss characteristic (Fig. 8) will be immune from the majority of drop-outs.

One suggested system uses out-of-contact recording and reproducing in a similar manner to a magnetic drum. The tape is run with its plastic backing in contact with a stationary drum, and the recording/reproducing head is placed with its gap facing the drum and separated by a small distance (say 0·000 5 in) from the magnetic-oxide layer. With such a non-contact system, oxide clumps and extraneous foreign particles either have no effect on the reproducing voltage or they increase it. Drop-outs are due only to thickness reductions in the magnetic layer or in the plastic base material. A study of Figs. 4 and 8 reveals that for n.r.z. recording the overall loss due to changes in base thickness is about 1·5 dB for a 0·000 1 in increase of separation beyond the normal 0·005 in, provided that a large recording current is used. Thus, if the total tape-thickness variation is the same as the base variation, a reduction of 0·000 4 in is the allowable tolerance for a maximum of 6 dB drop-outs.

The disadvantage of amplitude variations due to changes in base thickness may be overcome by using a double-head system,[2] where the stationary drum of the system just described is replaced by a second similar recording or reproducing head. The two heads are rigidly mounted so that the gaps face each other and are exactly parallel and opposite, as shown in Fig. 9. The
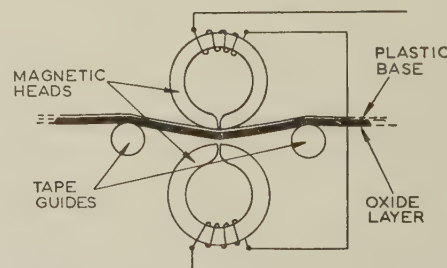
Fig. 9.—Diagram of two-head recording/reproducing system.

distance between the gaps is preferably equal to the thickness of the magnetic layer plus two thicknesses of base material, so that when the tape is transported with the base material in contact with one head the oxide layer is equidistant from the two. The variation of reproducing voltage in the two heads with movement of the recorded layer between the heads is shown in curves (a) and (c) of Fig. 8, where the abscissa represents the separation between the head surface and the adjacent oxide surface. The tape-base thickness is assumed to be 0·001 in. The two windings on the heads are connected in series, so that the induced voltages from changes of flux in the tape opposite the gaps are additive. The corresponding curve of the resultant reproducing voltage in the heads as a function of distance of the oxide from one of the heads is shown in curve (d) of Fig. 8.

Used in this way, loss of recording flux or reproducing voltage is avoided when a drop-out causes the tape to move away from one gap, since loss of signal by one head is compensated by gain in the other. For instance, on replaying an n.r.z. recording, the loss of signal for a displacement of 0·000 5 in from a single head is 6 dB, whereas a displacement of ±0·000 5 in with the two-head

stem produces a small increase of signal of about $1 \cdot 5$ dB. The ormal output from the two-head system is about 4 dB below at from contact reproduction on a single head for the conitions described.

(2.4) Accidental Pulse Recording

Having considered the factors which contribute to the loss of mplitude of pulse signals when recording and reproducing from ape, it is necessary to discuss the possibilities of recording and producing pulses which do not exist in the original signal. ccidental pulse recording can occur in n.r.z. recordings when a ape imperfection occurs in a position where no changes in cording current exist. If the tape were uniform in such a case, would be magnetized to saturation by the recording head in ne direction only. However, if the tape is momentarily lifted way from the recording head, the oxide layer may be lifted to a non-saturating field. Alternatively, if there is a small hole the oxide layer, a reduction of magnetization will occur and arface magnetic fields would be associated with such imperctions, as illustrated in Fig. 10. These local surface fields
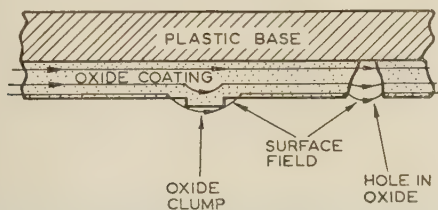
Fig. 10.—Surface magnetic fields near tape faults for d.c. recording.

ould give rise to voltages in the reproducing head which may be terpreted by the reproducing system as pulses in the recording gnal.

It is apparent that the causes of the noise pulses just described nd of drop-outs are due to similar tape imperfections. However, as will be explained, the relative severities of the two faults the reproducing signal are not the same for all types of imperctions, and separate testing methods are required to ensure nat a tape is perfect for data recording.

The change in reproducing voltage for both drop-outs and oise pulses depends on the rate of change of flux in the reproucing head, i.e. on the product of the tape speed and rate of nange of flux along the tape. For drop-outs, the latter figure determined by the rate of change of the recorded signal, which high compared with the rate of change due to the localized eparation loss. On the other hand, for noise pulses the rate of nange of the surface field depends on the rate of change of eparation loss. Hence, considering faults which cause separaon loss, the severity of drop-outs depends on the separation of ne tape from the heads, but the severity of noise pulses depends n the rate at which this separation occurs. It will be necessary, nen, in considering the design of a comprehensive drop-out test pparatus, to be able to detect both momentary reductions of the eproducing voltage of a pulse recording, and the spurious oltage pulses occurring on replaying a recording of a tape aturated in one longitudinal direction.

(3) DESIGN OF DROP-OUT TESTING APPARATUS
(3.1) General Design Considerations

can be concluded, from the considerations of the various cording and reproducing systems used for data storage on tape, na no single testing method could respond, by selection, to se particular drop-outs which occur in any one system. For asance, the commonest cause of drop-outs in the n.r.z. single-

head non-contact method (thin base) does not cause drop-outs with n.r.z. contact recording. Vice versa, separation loss, common with contact recording, does not cause drop-outs with the non-contact methods. Thus, contact and non-contact recording demand their own appropriate test methods, but since the majority of applications at present employ contact recording and reproducing this system is chosen for the test apparatus. The choice between n.r.z. and r.t.z. recording methods for the test apparatus is fairly easy, since either method can produce a comparable test for the other. N.R.Z. recording has the advantage for testing purposes that it is self-erasing and a tape may be automatically stopped when a fault occurs and recorded over the fault to investigate its permanence. The curves relating reproducing voltage output, head-to-tape separation and recording current (Figs. 4 and 7) for the two systems may be used to choose a recording current for n.r.z. recording which gives the same separation loss as an r.t.z. application. In this way equivalent n.r.z. tests can be performed when it is desired to test for r.t.z. applications.

The requirements for the test apparatus are as follows:

(a) To detect when any one of a number of reproducing signals is attenuated by a predetermined fraction of the normal maximum reproducing voltage.
(b) To register this event either as a permanent count or by stopping the tape machine.
(c) To register as a permanent count the number of reproducing pulses whose amplitude falls below the predetermined fraction.
(d) To be able to measure (a) and (b) over a range of frequencies, so that a number of different pulse densities and tape speeds may be accommodated.
(e) To detect and count reproducing voltage pulses of a predetermined amplitude occurring when unidirectionally saturated tape is replayed.

(3.2) Design and Circuit Arrangement of Drop-Out Tester
(3.2.1) Recording/Reproducing Equipment.

The equipment was initially designed to test tape in the range of 50–200 pulses per inch. N.R.Z. contact recording is used, and hence the maximum pulse rate records 400 reversals of magnetization per inch of tape, as shown in Fig. 2(b). Conventional ring-type recording and reproducing heads are used, with gaps to suit the pulse-density range and the type of recording, as discussed in Section 2. The recording head is slightly wider than the tape and the recording-head gap-length used is $0 \cdot 001$ in. A multi-track head is used for replaying the full track recording. Each track is $0 \cdot 045$ in wide and separated from adjacent tracks by $0 \cdot 015$ in. The reproducing-head gap-lengths are $0 \cdot 000\,5$ in. Most tests are made with tape speeds of 30 in/s and 15 in/s and a recording frequency of 3 kc/s, thereby recording the most commonly used pulse densities of 100 and 200 pulses per inch.

(3.2.2) Electronic Design.

A block diagram of the electronic functions is shown in Fig. 11. Each reproducing signal is fed into a monostable flip-flop, which delivers a square voltage output pulse each time the instantaneous input voltage exceeds a certain threshold level and so performs the function of discriminator and squarer. Between the reproducing head and this flip-flop a wide-band pre-amplifier produces a reproducing signal whose peak voltage can be set at a predetermined known multiple of the threshold voltage of the flip-flop. Hence, if any reproducing signal then reduces in amplitude to the threshold level, the constant-voltage square-wave output will stop. Fig. 12 shows the reproducing signals of two tracks, one having an amplitude reduction below the threshold level. The resulting absence of square-wave flip-flop pulses is also shown.

The square-wave flip-flop outputs are then combined in two gates. These gates are well known in computing circuits as
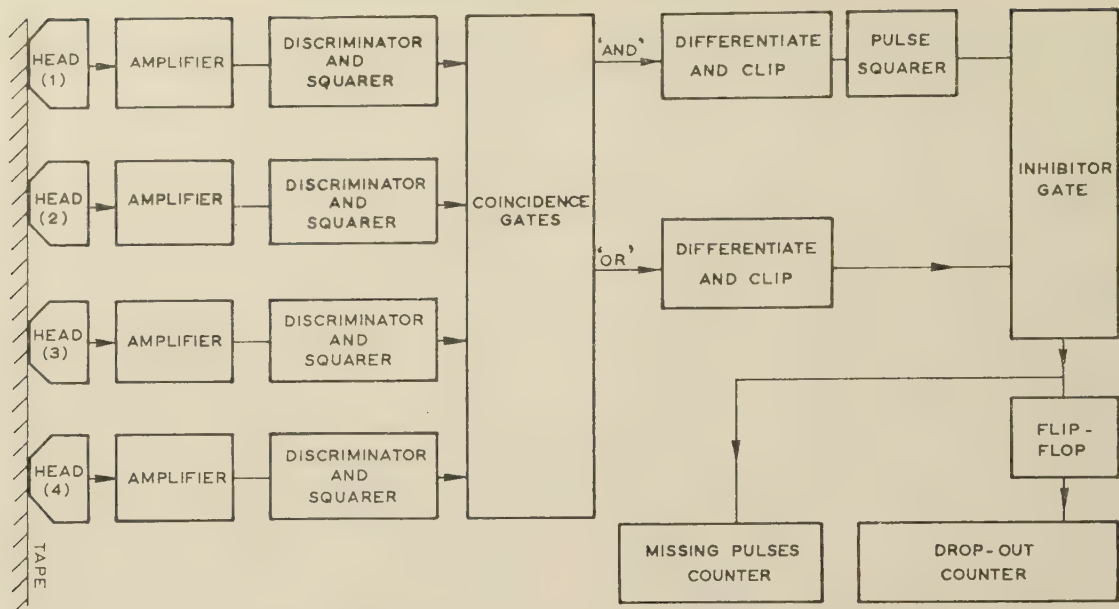
Fig. 11.—Block diagram of drop-out tester.

'and' gates and 'or' gates. The 'and' gate produces square-wave output pulses similar to the input pulses, provided that the pulses on all inputs are present. In other words, the instantaneous output voltage rises when all the inputs have risen, and it stays at this level until any one input falls. Hence, when a pulse is missing on any input the output also misses a pulse. Also, allowing for slight differences in phase of the input signals, the output voltage pulses will be slightly narrower than the input pulses. This narrowing of the 'and' pulses is a necessary feature of the system and is accentuated for clarity in Fig. 12(e). Usually, the greater the number of reproducing tracks the greater the overall phase difference between the flip-flop signals and the narrower the resulting 'and' pulse.

The second gate, known as an 'or' gate, is also connected across the flip-flop outputs, and this delivers a square-wave output if such a signal exists on any one of the inputs. In this case, as soon as any input voltage rises, the output of the gate rises similarly and falls again when all the inputs have fallen. The pulse output of the 'or' gate is therefore wider than the input pulses, if a phase difference exists, and it disappears only if all the inputs disappear [Fig. 12(f)].

Thus two signals are obtained, one which may be considered as a standard signal ('or') and one which disappears during a drop-out ('and'). Sharp pulses are derived from the 'or' signal [Fig. 12(h)] by using a differentiating circuit, and these are applied to a Dekatron counter which displays the total number of 'or' pulses applied to the counter. However, in the absence of a drop-out the 'or' pulses are prevented from reaching the counter by the 'and' signal. When the 'and' signal disappears the 'or' pulses are counted and hence the number of pulses in the drop-out is recorded. In order to perform the inhibiting function, the negative 'and' pulses [Fig. 12(g)] are fed to a monostable flip-flop which gives a positive square-wave for each 'and' pulse, of sufficient length to envelope the corresponding 'or' pulse [Fig. 12(j)]. These two signals are then added together and clipped to detect negative signals. These will occur only when the flip-flop output is zero at the time when an 'or' pulse occurs, i.e. when the 'and' signal is missing. Hence an output is obtained for the duration of the drop-out [Fig. 12(k)]. The 'or' pulses are also fed to a second counter via a long-time-constant flip-flop which passes a single pulse to the counter for each drop-out.
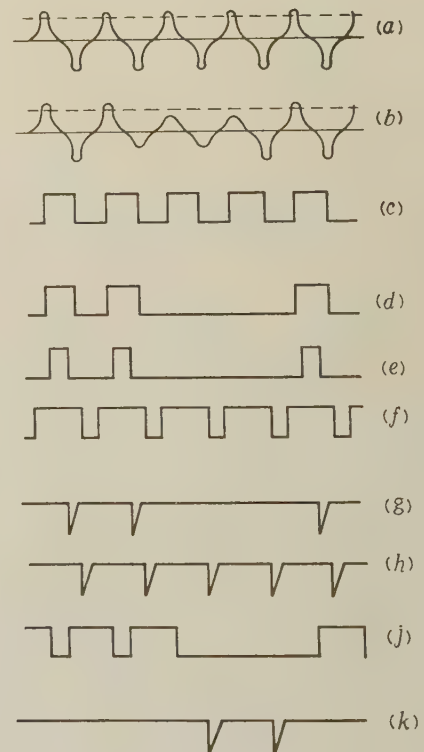
Fig. 12.—Waveforms in drop-out tester.

(a) Reproducing signal, head 1.
(b) Reproducing signal, head 2.
(c) Flip-flop output, head 1.
(d) Flip-flop output, head 2.
(e) 'And' gate output.
(f) 'Or' gate output.
(g) Differentiated 'and' negative signal.
(h) Differentiated 'or' negative signal.
(j) Monostable flip-flop triggered by (f).
(k) Added (h) and (j) negative signal.

(.2.3) Detection of Noise Pulses.

In addition to drop-out detection and counting, a second function of the test equipment is the ability to count spurious producing voltage pulses which may occur on a unidirectionally saturated tape. The recording signal in this case is a constant recording field produced by a d.c. head current equal to the normal peak head current used in pulse recording. In the reproducing chain, the gain of the pre-amplifier is increased so that the output reaches the threshold voltage of the monostable flip-flop for an input signal which is one-tenth of the normal output of a pulse recording. Thus, a noise pulse of this amplitude occurring in any one of the reproducing channels will operate the appropriate flip-flop, and an 'or' signal will be generated. No 'and' signal will occur until all the channels are simultaneously operated by the noise signal, and this is a most improbable event. Hence, the 'or' signal will be counted, and in this way the noise pulses occurring on replaying a saturated tape are recorded.

(4) TAPE PERFORMANCE IN DATA RECORDING
(4.1) Drop-Outs

The technique of examining signals recorded on tape by making the recording visible with colloidal magnetite finds useful application when the physical cause of a drop-out is not immediately apparent. In Fig. 3 two faults are shown in the oxide layer, each of which is only about 0·005 in long. The faults are small oxide clumps which would cause separation loss on reproduction. A similar but larger fault is shown in Fig. 13(a), and its effect on the recorded signal can be seen in Fig. 13(b). Here, a large area of tape has been lifted away from the recording head by the fault, with consequent blurring of the colloidal magnetite deposits. Where separation loss has occurred, the surface fields of the recording are more diffuse longitudinally, owing to spread of the transition region, as shown by broadening of the lines in Fig. 13(b). The heavy black lines on the fault indicate strong magnetization where the clump has made good contact with the recording head.
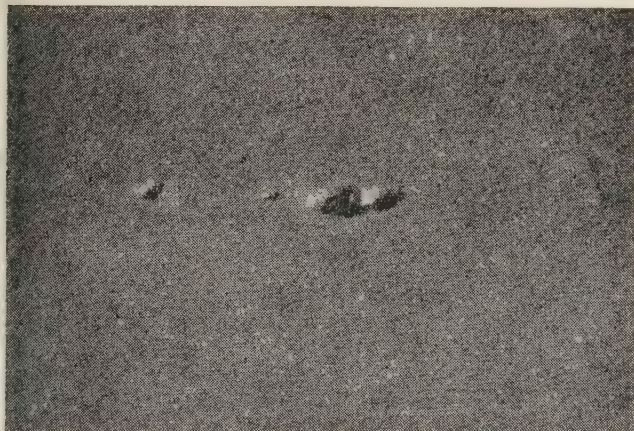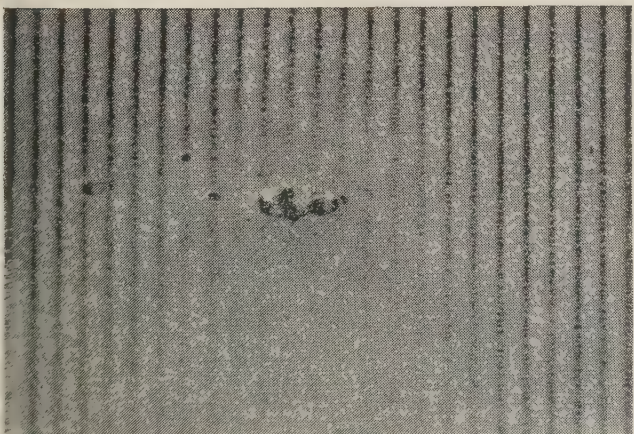
An examination of drop-outs occurring on samples of commercially available tapes has shown that, provided that sufficient care is given to the production of a uniform magnetic lacquer with no foreign particles and that the tape is manufactured under dust-free conditions at all stages, those having cellulose-acetate or polyester bases are superior to others. Polyvinyl-chloride bases of Continental manufacture are used extensively in this country for audio-recording-tape bases, but, in general, the quality is poor for data-recording tapes, owing to the high incidence of local thickness changes. Small batches of high-quality p.v.c. have proved to be excellent as base material, but such quality is available only in sample lengths. Cellulose acetate and, more recently, polyester are used at present for data-recording-tape bases. Both materials are manufactured with a high degree of uniformity and with virtual freedom from surface blemishes. The inherently low susceptibility of polyester to moisture absorption and its resistance to tearing have made it more favourable for data-recording tape.

All the tapes tested in this investigation have similar magnetic properties and consist of acicular particles of γFe_2O_3 embedded in a flexible binder suited to the various base materials. The particles are about $1\,\mu$ long and have a length/width ratio of 5 : 1 to 10 : 1. The anisotropy of particle shape is the controlling factor in the magnetic behaviour of the tape, and in most modern tapes the particles are oriented along its length, resulting in a high ratio between remanent and saturation inductions for fields applied in this direction. Fig. 14 shows a family of hysteresis

(a)

(b)

Fig. 13.—Effect of drop-out on recorded pattern.
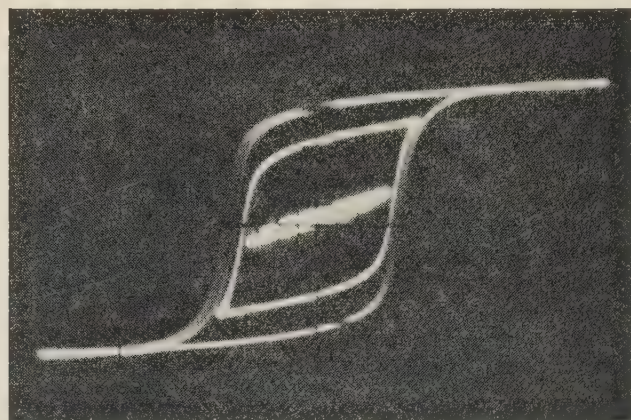
Fig. 14.—Hysteresis curves for magnetic tape.

curves, for a tape of the type described, taken from a cathode-ray tube 50 c/s B/H plotter. The coercivity is about 260 oersteds and the maximum remanent flux per 0·25 in tape is about 0·7 webers.

Drop-out tests have been performed on 0·5 in wide cellulose-acetate-based tapes and polyester tapes, using a pulse density on the tape of 200 per inch per track. Eight tracks have been tested at a drop-out level of 8 dB. Tapes made of cellulose acetate of British manufacture have an average length of about 400 ft between drop-outs; for American acetate this length is 700 ft, and for recent polyester, 1 100 ft. The figures quoted refer to

magnetic layers which have been manufactured under similar conditions, and so they reflect to some extent the quality of the base material. Examination of some American tapes, which are claimed to be manufactured for data recording, shows that nearly all the drop-outs are due to faults which momentarily lift the tape from the recording or reproducing head. At the recorded pulse density of 200 per inch and a drop-out level of 6 dB, about half of the detected drop-outs are caused by removable defects, such as chips of base material. These were either not cleaned off during the slitting process or were shaved off by the guides of the tape deck. Most of the permanent faults are also due to base defects, such as dents or small air bubbles. Clearly, in this case, the tape manufacturer cannot further improve his product apart from exercising caution with regard to foreign particles. Future progress depends more on the availability of flaw-free base material and improved systems for handling the tape. In the latter case, it seems that the out-of-contact systems described might be used with advantage.

(4.2) Noise Pulses

In the discussion of the causes of accidental pulse recording and reproduction with an n.r.z. system, it was concluded that, although the same physical tape faults cause both drop-outs and noise pulses, the relative severities are not the same for all causes. The drop-out test apparatus has been used to count reproducing voltage pulses of a known fractional amplitude of the normal reproducing output from a pulse recording. At given noise and drop-out levels it is found that small nodules on the surface of the magnetic layer cause more noise pulses relative to drop-outs than do creases in the tape or thickness variations of the magnetic layer. These results bear out the suggestion that drop-out tests and noise tests are not equivalent, and that sharp surface irregularities must be avoided to minimize noise pulses.

(4.3) Conclusions

It is of interest to formulate some idea of the limiting resolution of present-day tapes for recording and reproducing conditions

similar to those used in the test method. An essential safeguard in data recording is that adequate signal voltage should be obtained, so that electronic operations on the replayed data may be performed without the hazard of electronic equipment noise being comparable with the signal. Upper and lower discrimination limits are set on the reproducing signal amplitude, so that no signal amplitude variations reduce the signal to the upper discrimination limit and no noise pulses reach the lower limit. The maximum pulse density which can economically be used under these conditions is about 200 per inch of tape for a track width of 0·045 in. If the discrimination levels are set for equal probabilities of occurrence of drop-outs and noise pulses, the largest voltage difference between them is obtained for minimum overall probability of fault occurrence.

Drop-out tests and noise tests indicate that at a pulse density of 200 per inch, the probabilities of occurrence of drop-outs and noise pulses are equal at about 50% and 5%, respectively, of normal maximum output for a track width of 0·045 in. However, at present the probability of occurrence of faults at these levels is rather high, resulting in a low yield of fault-free tape. Some relaxation of the limits would seem to be economical for existing British tapes, although the figures given may be taken as a short-term target for n.r.z. contact recording systems using tape speeds of about 100 in/s.

(6) REFERENCES

(1) HOAGLAND, A.: 'Magnetic Data Recording Theory: Head Design', *Communication and Electronics*, 1956, **27**, p. 506.
(2) MEE, C. D.: British Patent Application No. 24583: 1955.

DISCUSSION BEFORE THE MEASUREMENT AND CONTROL SECTION, 18TH FEBRUARY, 1958

Mr. D. W. Willis: The paper confines its attention to drop-outs and noise pulses (or drop-ins), and does not mention other parameters of importance in data recording, such as tolerance on slitting, the consistency of the thickness of the backing, behaviour under variable climatic conditions, the quality of the binder, etc. These factors are of vital importance to designers of tape systems and should be considered by the authors.

The designer of a tape system must from the start recognize the incidence of drop-outs, and that they are of two kinds. There are permanent drop-outs, which are located in fixed positions on the tape, and transient drop-outs, often caused by spacing losses due to surface imperfections, dust, loose oxide or shavings off the backing, which are often non-repetitive and not located in fixed positions. The optimum design of any particular system is conditioned largely by the distribution in length and separation of both types of drop-out.

For example, if the designer knew that all drop-outs were shorter than some maximum length, he could allocate to each digit a length of tape twice as long and repeat the digit at high density throughout the space, integrating the information on reproduction over the whole of this space. The tape could then be used without any other special provision for drop-outs. Alternatively, the same information could be recorded two or three times, an error-detecting code, or the identity of two out of three, being taken to show which reading was correct. Which of these two methods is chosen depends on the distribution of the drop-outs.

Another method more widely used is to record at the extreme density the tape will give, and to remove permanent drop-outs either by subjecting the tape to a severe selection test, or by addressing the tape only where it should be used, or by marking the tape in some way when it should not be used. Transient drop-outs are dealt with by recording a block and by immediately playing it back and checking by some arithmetic, parity or identity check. If the information is correctly played back, it is assumed that at some subsequent time it can again be played back correctly, providing that one or two attempts are made.

The non-contact method has an entirely different approach. Here the performance of the tape is purposely derated, and a better and more uniform performance is obtained from the system. Fig. A shows the sort of drop-out performance that can be obtained from out-of-contact recordings made on a commercially available equipment. As can be seen, there are no drop-outs recorded below about 50% of the normal signal level. Non-contact methods have other advantages. The author mentions one of them—that wear is minimized. More important is the fact that the maintenance required by the heads is minimized. In some contact systems, the heads have to be cleaned on almost every run, but in non-contact operation far less cleaning is necessary.

The method of getting prints off tapes is a useful way of showing up faults. My colleagues have used a similar method to check track alignment and to give a very accurate measurement of tape speed.
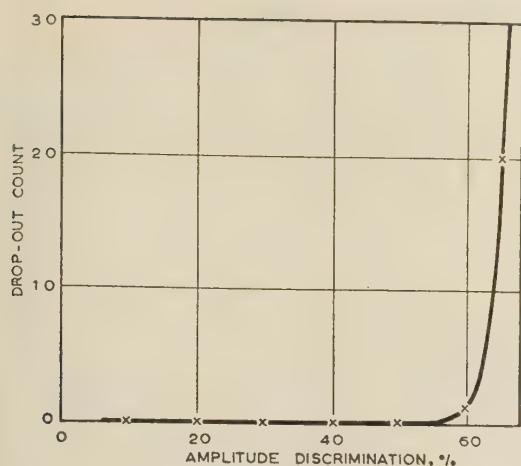
Fig. A.—Drop-out distribution for out-of-contact recording.
3 600 ft of tape.

The problem of specifying the drop-out performance of tape is receiving close consideration on several committees at the present time. One method proposed is to measure drop-outs on a narrower track width than anyone uses, and to test all the tape in this manner. Another is that the tape should be fully saturated, then run through an equipment with high amplifier gain. Imperfections in the tape should show up as non-white noise. The author points out that imperfections do not affect the drop-out and drop-in counts equally, so I assume that such a system could not immediately be useful as a simple means of measuring the uniformity and performance of tape. I would be glad if the author would comment on this point.

He has supplied some information which is useful to the system designer, but much more information, particularly in respect of the statistical distribution in length and separation of drop-outs, is required. It is only the supplier who can conveniently provide this information.

Dr. J. E. R. Young: I should like to make a few remarks about testing data tape.

The data storage system uses ½ in tapes with seven recording tracks 0·033 in wide. The design of the digital system is such that no drop-outs and no spurious pulses (drop-ins) can be tolerated, and all tape is to be interchangeable. Therefore, every track of every reel is tested.

The tape tester was constructed to resemble the tape-handling system as completely as possible. The test sequence is: saturate record with a continuous sequence of pulses; test that all reproduced pulses top an upper limit; erase; check that no noise pulse exceeds the acceptable lower limit; all this is done in a single pass of the tape over a set of heads. When a defect is recognized the tape is stopped and the offending portion examined through a microscope, and very frequently the women operating these testers can remove the defect by 'surgery' using a scalpel, or by 'dry cleaning' with solvents and cotton wool. After that the machine is restarted; it back-spaces and tests that portion of the tape and continues if the repair is satisfactory. When a defect will not clean up, that reel is returned to the manufacturers, but when a completely tested reel is obtained it is packaged and sold as guaranteed error-free tape.

This does not completely cure the tape error problem, as errors may still arise after long use, and in the future, when dealing with higher-density recording and thinner tape, a different

philosophy is envisaged. For example, a 'two-gap head' has been developed with a recording and a reproducing head closely adjacent and fixed in the same assembly; with this the data recording machine can check everything it has recorded and if this is not correct it can erase that portion of the record and try again further along the tape. But it must be emphasized that this device, while it does greatly improve reliability, should not be used with low-grade tape, for it is cheaper to use expensive pre-tested tape than to hold up a large computer while groping for an error-free portion of tape.

Looking at Fig. 13 of the paper, I wish to point out that a soft and resilient pressure pad could have made the tape conform much better to the recording head, and the total amount of signal lost would have been much less.

Finally, it seems, from the paper and this discussion, that we should consider using a special type of tape for pulse recording instead of the high-quality audio tape which is used nowadays.

Mr. G. M. E. Williams: I am prompted to ask how much more reliable magnetic tape has to be. Some intending users of computer systems for commercial data with magnetic tape are looking for an occurrence of errors of not greater than about 1 in 10^7. The paper shows that drop-outs occur about 1 in 1 100 ft of some polyester-based tape. At 100 digits to the inch, the improvements in magnetic tape in respect of occurrence of drop-outs would therefore have to be about eight times greater not to exceed 1 in 10^7. How soon can such an improvement be achieved?

Mr. Willis referred to making two or three recordings in parallel at a time and comparing one against another to guard against drop-outs. I support his statistical approach to the problems of magnetic tape, but if he looks at the statistics of the use of two or three parallel channels he may find that the gain in reliability compared with using a single channel does not justify the additional equipment capacity and time involved.

Mr. F. J. M. Laver: The author says that in a magnetic tape data-recording system the signal 'must be reproduced without any loss or addition of pulses', but this is an unattainable ideal, and data-handling systems should be designed on an economic basis to meet a certain probability of error, because errors are inevitable. Any design which requires perfect tapes is immature.

The results given in Section 4.1 show average lengths of 400 ft or more between drop-outs, but it is not clear whether this is for a drop-out on any of the eight tracks, or is the separation of drop-outs on individual tracks. If the former, this at 200 pulses per inch corresponds to a drop-out rate approaching 1 in 8×10^6. Communication engineers would be happy to have a system with so low a probability of error.

Fig. 5 shows return-to-zero recording which uses only two levels of magnetization—positive and zero, but above Fig. 7 there is a reference to three states of magnetization that have to be distinguished. I am not clear whether this is meant to apply to a 3-level r.t.z. system. Such a system resembles a non-return-to-zero one, but zero magnetization could indicate that no signal is recorded. Will the author confirm that the references to r.t.z. recording in the paper are to 2-level systems?

It would be interesting to know whether the author has any statistics showing how the proportion of drop-outs, or the average length between drop-outs, varies according to the width of the track used for the recording? For drop-out tests the author records across the whole width of the tape: this is not what will be done by a user. Is the author satisfied that this does not significantly affect the number of drop-outs found?

[The author's reply to the above discussion will be found overleaf.]

THE AUTHOR'S REPLY TO THE ABOVE DISCUSSION

Dr. C. D. Mee (*in reply*): First, I would like to thank Mr. Doust for reading the paper in my absence and replying at the meeting.

The tape manufacturer is constantly improving the quality of recording tape with respect to the factors mentioned by Mr. Willis. It is felt that current techniques for tape production are entirely adequate to meet the needs of the data-tape user as regards width and thickness tolerances. The quality of the binder material has been greatly improved recently in order to satisfy the severe conditions obtaining in video tape recorders. Here, relative head-to-tape speeds of 150 ft/sec may be encountered, and present-day tapes are capable of recording and replaying at least 100 times. The wearing qualities of tape for data recording have also been improved recently by applying a thin plastic layer over the oxide coating.

The drop-out counts given in Section 4.1 refer to the error-free distances when eight tracks are simultaneously examined. It should be added that the figures quoted were obtained at least a year ago on experimental tapes coated under similar conditions. The large difference in drop-out counts for the different bases underlines the importance of the quality of the base material. In general, tapes have now been improved to the stage where freedom from 6 dB drop-outs at 200 pulses per inch, with non-return-to zero recording on 0·045 in wide tracks, may be considered a reasonable commercial standard.

Some progress towards a standard test procedure for data recording tape has been made, owing to the growing popularity of four tracks per quarter-inch of tape with 0·06 in between track centres. Standardization of the track width at 0·03 in seems likely and, in this event, the area of tape to be tested for drop-outs can be specified. The relation between noise pulses from a saturated tape and the drop-out count is not considered sufficiently good to use noise testing instead of drop-out testing. It is suggested, then, that a realistic specification for data recording tape for the next few years arises from the track widths, pulse densities and discrimination levels quoted here. At the present, the yield of 1 200 ft lengths of drop-out-free tape under these test conditions should be high. The availability of large quantities of high-grade polyester base should quickly alleviate the limitation of 1 200 ft for high yield.

The return-to-zero recording analysis described in Section 2.2.2 refers to the two-level system illustrated in Fig. 5. The three-level system was mentioned with respect to its susceptibility to longitudinal pulse spreading in recording. The width of the recorded track appears to have a negligible effect on the drop-out count when comparing drop-outs from the full track recording of the test apparatus with a multi-track recording made with a head similar to the reproducing head.

PHASE-COHERENT BACK-SCATTER OF RADIO WAVES AT THE SURFACE OF THE SEA

By E. SOFAER, Associate Member.

(The paper was first received 11th December, 1957, and in revised form 21st January, 1958.)

SUMMARY

Soon after the completion of the B.B.C.'s television transmitting station in Devon and the establishment of a full-scale service, complaints were received from coastal areas around Plymouth that the transmission was subject to rhythmic variations in amplitude. The investigations which followed these reports are here described. The variations are found to be due to phase-coherent back-scatter from the sea, and to depend on the configuration of the surface of the sea. The phenomenon is examined theoretically.

LIST OF SYMBOLS

v = Velocity of propagation of sea waves.
λ_s = Wavelength of sea waves.
g = Acceleration due to gravity.
λ = Wavelength of television transmission.
f_s = The isostematic radio frequency.
f_1 = A radio frequency not isostematic.
n = Number of wavecrests over back-scattering surface of the sea.
P_t = Transmitted power.
P_r = Received power.
σ = Back-scattering cross-section.
$k_1, k_2 \ldots k_6$ = Constants of proportionality.
r = Distance of back-scattering surface from receiving aerial.
A = Physical area of back-scattering surface.
ρ = Electrical properties of sea water.
s = Statistical distribution of elementary scattering surfaces.
r_h = Distance of radio horizon from receiving aerial.
r_e = Distance of the edge of the cliff's shadow from the receiving aerial.
r_0 = R.M.S. distance between r_h and r_e.
S = Barometric tendency when steady.
U = Barometric tendency when unsteady.

(1) INTRODUCTION

The British Broadcasting Corporation inaugurated a television service in Devon in December, 1954, with a temporary low-power installation and a radiating system mounted 150 ft above ground level. The site of the transmitter, a peak on Dartmoor known as North Hessary Tor, is 1 675 ft above sea level and about 15 miles inland from Plymouth Sound.

The station was completed and the present full-scale service established in May, 1956. Besides the increase in the effective radiated power which occurred at the time of the change, the height of the radiating system was increased to 633 ft above ground level. The transmissions were vertically polarized and the frequencies radiated, in both the temporary and final conditions, were 48·25 and 51·75 Mc/s, the former carrying the sound information and the latter the vision.

A few days after the full-scale service had been established, reports were received that the transmission was subject to interference of an unusual nature in coastal areas between Plymouth and Looe. This took the form of rhythmic variations in signal amplitude accompanied by a series of echoes which stretched across the viewing screen. The period of the variations was about 40 cycles per minute. The places where the complaints originated were shaded from the direct transmission by cliffs and had a clear view of the sea.

A preliminary investigation of these reports was made with routine field-strength-measuring equipment and simultaneous records of the amplitude variations at both the sound and vision frequencies were taken on a recording milliammeter. The interference was also observed on a television receiver. While these tests did little more than confirm the reports, they did point to a strong connection between the effect and the proximity of the sea. It became evident that for a better understanding of the phenomenon more elaborate equipment would be necessary, and a further investigation was therefore planned in which, with the aid of directional aerials and a waveform monitor incorporating a time calibration, it was hoped to discover the source of the offending echoes.

(2) EQUIPMENT

The arrangement of the equipment used is shown in schematic form in Fig. 1. The signal to be examined was picked up on a
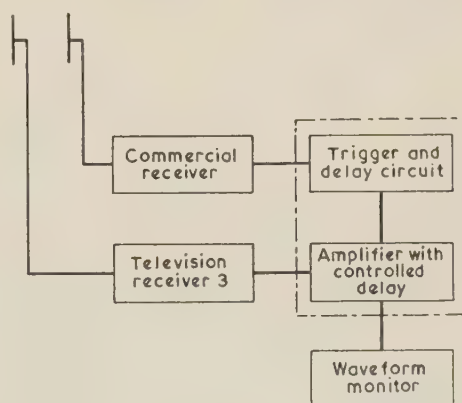
Fig. 1.—Arrangement of the equipment.

directional aerial capable of being raised to any height up to 36 ft above ground level and rotatable in any direction. This aerial was connected to a receiver of B.B.C. design (TV/REC/3), the output of which was taken through a variable-delay amplifier to a waveform monitor. The triggering supplied to the variable-delay circuit was obtained from the line-synchronizing pulses of a commercial receiver connected to a separate aerial. This aerial was maintained constantly in the direction of minimum interference. The purpose of this arrangement was to ensure that the amplitude of the trigger pulse supplied to the waveform

monitor would not vary with rotation of the search aerial, and that the display would be held at a constant and convenient part of the viewing tube on the monitor.

The search aerial had a horizontal radiation pattern of cardioid form with a front/back ratio of 19 dB. The second aerial was a commercial H-type dipole and reflector with a front/back ratio of 12 dB.

(3) OBSERVATIONS

One of the areas most affected by the phenomenon is Kingsand, on the western side of Plymouth Sound, but the observations made there were confused by the presence of reflecting objects in the Sound and on the opposite shore. The breakwater, with its lighthouse at the western end, and the masts of the Air Ministry Communications Station at Staddon Heights produced major echoes which, together with the intervening stretch of sea, were at first thought to be connected with the variations. Subsequent tests at Downderry, a village between Plymouth and Looe, on an open site remote from reflecting objects, gave a clearer picture of the nature of the phenomenon. Tests were also made at Staddon Heights, and, using a portable television receiver, in a boat in the Sound. The locations of these sites

frequencies of the two channels, and to establish that the interference was originating at a common moving reflector.

Sections of the records are reproduced in Figs. 3(a) and 3(b). From the time and relative-field-strength scales shown, the range and frequency of the variations at each frequency may be seen. A third scale giving a count of the number of cycles recorded over a length of chart is also shown. It is clear from a comparison of each pair of simultaneous records that no exact relationship exists between the variations at the sound and vision frequencies.

(3.2) Method of Observations

The operational procedure adopted with the waveform monitor was to observe the line-synchronizing pulse while picture information was being transmitted. Most of the work was done in this way and the account given in the present paper refers generally to these observations. Between scheduled transmission times, a 2 microsec positive pulse was radiated especially for these tests, and was useful for identifying permanent echoes.

It was not possible to record photographically the displays on the monitor. The exposures required by this equipment to obtain a satisfactory negative are of the order of 10 sec, but the

Fig. 2.—Locations of test sites.

are shown in Fig. 2, and details of their heights and distances from the sea recorded in Section 12.1.

(3.1) Amplitude Records

Simultaneous recordings of amplitude variations at the sound and vision frequencies were made at Kingsand during the preliminary tests. The site used is the more southerly of the two shown in Fig. 2. It was hoped by these to find that the interference frequencies were in the same ratio as the carrier

signal traces varied so rapidly that an exposure greater than $\frac{1}{25}$ sec would have produced a useless record. In consequence, the forms taken up by the trace are drawn free-hand and shown in Figs. 4–7.

(3.3) Observations at Downderry

Fig. 4 is the standard form of a complete line. When the amplitude variations were small, the trace of the line-synchronizing pulse obtained with the aerial maximum directed towards the transmitter resembled the standard form fairly

Fig. 3.—Comparison of the amplitude variations at the sound and vision frequencies (48·25 and 51·75 Mc/s respectively).

(a) Demonstrating the absence of an exact relationship between the beats on
(b) the two frequencies.
(c) An example of alternation in the amplitude of the variations between the two frequencies.
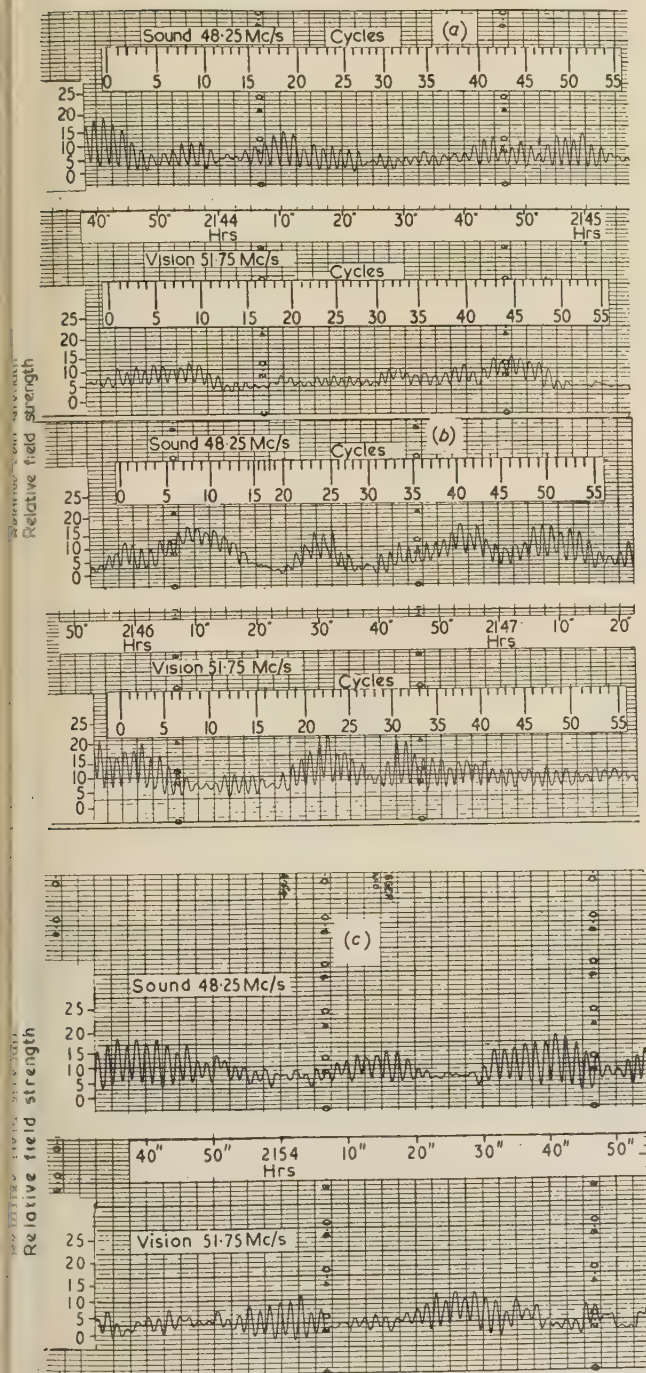
Fig. 4.—Standard form of a single line of picture transmission.

closely. The portion representing black level, which should have been straight, was distorted by small ripples, and a ripple of similar width appeared at the bottom left-hand corner of the line-synchronizing pulse, as shown in Fig. 5(a). At times of

Fig. 5.—Typical form of the distortion of the line-synchronizing pulse at Downderry.

(a) Aerial maximum towards the transmitter; moderate interference.
(b) Aerial maximum towards the transmitter; strong interference.
(c) Aerial maximum towards the sea; strong interference.

strong interference the trace was of the form shown in Fig. 5(b). It was particularly noticed that at such times the irregularities were greatest at the commencement of the line. This peculiarity is discussed in Section 4.3.

When the aerial maximum was oriented seaward in any direction over an arc of about 200°, distortion of the waveform was a maximum, even to the extent that the line-synchronizing pulse disappeared completely. Fig. 5(c) suggests the form of the trace for this condition.

(3.4) Observations at Kingsand

The observations on the waveform monitor were made at the more northerly of the two sites shown in Fig. 2.

Distortion of the waveform is greater at Kingsand than at Downderry, even in the absence of interference. Two permanent echoes lie inside the line-synchronizing pulse, reducing its effectiveness for triggering. They appear to originate at the Air Ministry masts at Staddon Heights and the lighthouse on the breakwater. When interference is present additional echoes further destroy the waveform. Fig. 6(a) represents the type of
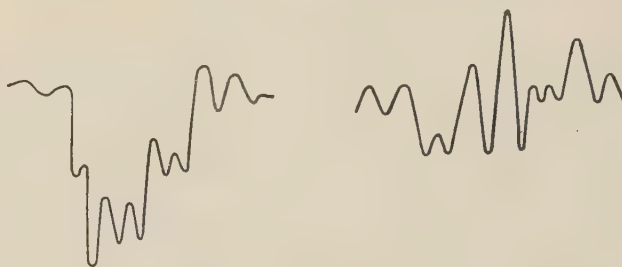
Fig. 6.—Typical form of the distortion of the line-synchronizing pulse at Kingsand.

(a) Aerial maximum on a bearing mid-way between the transmitter and the masts at Staddon Heights; fairly strong interference.
(b) Aerial maximum towards the sea; strong interference.

trace obtained with the aerial maximum midway between the direction of the masts and of North Hessary Tor. In seaward directions, and when the interference was strong, the trace took the form shown in Fig. 6(b).

(3.5) Observations at Staddon Heights

Measurements were made at a site 350 ft above sea level on a line optical to the transmitter in the one direction, and with a clear view of the sea in the other. The field strength at this site was estimated from the deflection of the a.g.c. meter to be 3 mV/m. With the aerial maximum towards the transmitter and the minimum to the sea, a nearly steady trace was obtained, the tendency to vary being only just perceptible. With the aerial minimum to the transmitter and the maximum to the sea, distortion of the waveform was significant. This condition is depicted in Fig. 7.

15

Fig. 7.—Typical form of the distortion of the line-synchronizing pulse at Staddon Heights.

Aerial maximum towards the sea.

Voltage measurements were made of the amplitude of the distortion. The result, after correcting for aerial directivity, showed the magnitude of the reflected signal to be 7% of the direct signal from North Hessary Tor. Its strength is therefore estimated as 200 μV/m, comparable with the mean field strength at Kingsand and Downderry.

(3.6) Observations on Plymouth Sound

A portable television receiver was installed in a boat, and a tour was made round the breakwater in Plymouth Sound, starting at the Royal Dockyard.

An arc of 15 miles radius centred on North Hessary Tor marks the approximate edge of the shadow cast in Plymouth Sound by the intervening land. Between the Royal Dockyard and this line, the amplitude variations were present continuously. When the line was crossed and the receiving aerial was directly irradiated by the transmitter, the variations were no longer apparent. They reappeared, however, each time the aerial was shielded from the direct radiation, e.g. by a destroyer anchored in the Sound, by Drake's Island or by Staddon Heights, indicating that the interference is always present, even when masked by the strong field of the direct radiation.

(3.7) Observations on Band II

Test transmissions in Band II on horizontal polarization were being radiated from North Hessary Tor at the time when the investigations which form the subject of the paper were being made. Advantage was taken of the opportunity to observe the Band II signal in areas where the television signal was susceptible to interference. An f.m. receiver was taken to Downderry and listening tests were made. Interference was present and, allowing for amplitude limiting in the receiver, appeared to be equal in severity to that experienced in Band I. No reliable estimate of its dependence on polarization can, however, be made from this observation. It was accompanied by multipath distortion.

It became clear from this experience that the effect could occur over a wide range of frequencies. That it was related to the sea clutter observed by operators since the earliest days of radar seemed evident, but the first report of rhythmic variations appears to have come from Crombie,[1] who observed them at a frequency of 13·56 Mc/s.

(4) MECHANISM OF THE PHENOMENON

Crombie has explained the beats in terms of a Doppler shift in the frequency of the transmission. The effects observed near Plymouth are considered from this point of view.

The combined evidence of the preliminary and present tests is:

(a) That the echoes come directly from the sea.
(b) That the phenomenon occurs over a wide range of frequencies.
(c) That the pattern of interference is different at different frequencies.
(d) That it occurs on both horizontal and vertical polarization.

(e) That it is present at sites directly irradiated by the transmitter as well as at sites shielded from the direct transmission; on high ground or at sea level, provided that the receiving aerial is not very distant from the sea.

The surface of the sea may be described as diastematic. This unusual word, which is absent from most dictionaries, is defined in the Oxford English Dictionary as 'characterized by intervals'. From this may be derived 'isostematic', i.e. characterized by similar intervals. Isostematism therefore bears the same relation to distance as synchronism bears to time.

The scatter surfaces on the sea are the sides of the waves which face the transmitter. Let the spacing between them be such that the phase angle at the receiving aerial between rays scattered back at adjacent surfaces is 2π, as illustrated in Fig. 8. The
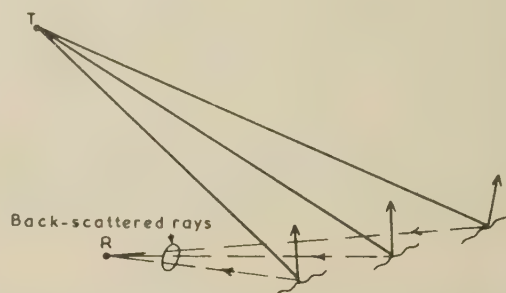
Fig. 8.—Diagrammatic representation of the rays scattered back from the sea.

The phase difference at R between adjacent rays is 2π.

rays will add directly, and a signal of greatly increased amplitude will reach the receiving aerial. Such spacing may be said to be isostematic with respect to the transmitted wavelength.

An isostematic system of sea waves, taking a normal course over the surface of the sea, retains its isostematic character for periods of time which are long compared with one period of their motion. When such a system travels towards or away from the receiving aerial, the phase of the back-scattered energy changes cyclically at a rate depending on the velocity of the motion. In the presence of the constant direct transmission these changes in phase produce a periodic variation in the amplitude of the received signal. If the amplitude of the back-scattered energy is comparable with that of the direct transmission, large variations occur, a situation well illustrated by the recordings reproduced in Fig. 3.

(4.1) Correlation with the Observed Beat Frequency

The relationship between the velocity of the sea wave and its wavelength λ_s is given by the expression

$$v = \sqrt{\left(\frac{g\lambda_s}{2\pi}\right)} \qquad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad (1)$$

where g is the acceleration due to gravity. This formula is accurate for waves of small amplitude occurring in deep water. The phase difference between waves radiated from a single source and arriving at a point in space by separate paths is $2n\pi$ radians (n integral) when their path differences are multiples of λ, the transmitted wavelength. In the first instance, let it be assumed that the direction of back-scattering is in the plane of propagation. If the back-scattered rays all arrive at the receiving aerial in phase, the reflecting surfaces must be spaced $\frac{1}{2}\lambda$ apart, i.e. $\lambda_s = \frac{1}{2}\lambda$. At the vision frequency, $\lambda = 5\cdot8$ m and $\lambda_s = 2\cdot9$ m, so that $v = 2\cdot13$ m/sec.

One cycle of change occurs when the wavecrests travel a distance 2·9 m, and the periodicity of the change is therefore 0·73 c/s. The number of cycles recorded in Fig. 3(a) during the

Fig. 9.—Elliptical arcs based on North Hessary Tor and Kingsand, indicating the directions of the sea waves which will produce interference at Kingsand.

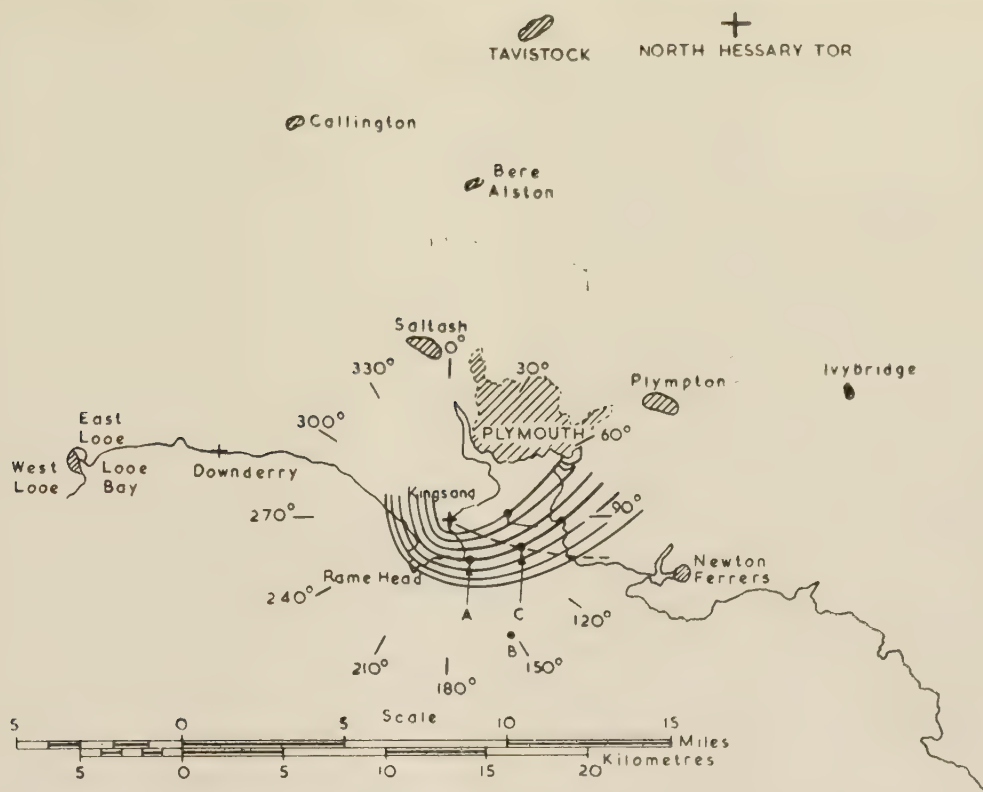inute 2144–2145 hours G.M.T., and in Fig. 3(b) during the inute 2146–2147 hours G.M.T. is in each case about 40·5, a eriodicity of 0·68 c/s. The calculated periodicity is therefore o great by 8%.

In Fig. 9, elliptical arcs are drawn whose foci are at North essary Tor and at Kingsand, at the site where the records were ken. For the energy incident on the sea to be diverted to ingsand, the wavecrests must run tangentially to the arcs. ays touching the sea at points on any one arc arrive at ingsand with equal delays.

The arcs are not equidistant from each other along their ngth, but diverge as they move to either side of their common ial direction. The value 0·5λ is applicable to λ_s only where e reflecting surfaces are normal to the common axis. Else- here, λ_s is greater. For the site used, reflection must occur st of the line on which the positions A and B marked in g. 9 lie, because for all that part of the sea to the west of this ne the receiving aerial is shielded by high ground. The values λ_s at A and B are 0·56λ and 0·55λ, respectively, as calculated Section 12.2, and the periodicities of the beats derived from em are, in the one case, 0·69 c/s and in the other, 0·70 c/s. o corrections have been applied in the foregoing calculations allow for the limited accuracy of the formula used.

it should be pointed out here that a change in the position of e back-scattering area requires both the wavelength and the rection of the sea waves to change appropriately, if isostematism be maintained. The term 'isostematism' therefore connotes re than a relationship between lengths, namely the wave- ths of the radio transmission and of the sea waves. It also des the concept of direction, implying that the wavefront the sea waves is suitably oriented.

In Fig. 9, the position C has been so chosen that the periodicity of the interference reflected at this point agrees exactly with the observed value of 0·68 c/s after correcting for the amplitude of the sea wave. Eqn. (1) yields a result for v which is too small[2] by an error roughly of magnitude h/λ_s, where h is the crest-to-trough amplitude of the wave. In deep water, h/λ_s has a maximum value of 0·14. The correction applied in the present case is 0·1.

The depth of the Sound at C is very much greater than $\frac{1}{4}\lambda_s$, and no correction for depth is therefore necessary. The position of C marks one of the possible areas where reflection could have occurred during recording. The locus of such areas is the dashed line through C. Its direction accords well with observations. Interference was maximum when the aerial beam was turned seaward along a line tangential to the coast south of Newton Ferrers.

It would appear at first sight that, because of the Doppler mechanism, the frequency ratio between the beats present on the two radio transmissions should be the same as that between the carriers. The recordings taken during the preliminary tests and reproduced in Figs. 3(a) and 3(b) establish that no such relationship exists between the frequencies of the beats on the two channels.

It has been seen, with the aid of Fig. 9 and the calculations in Section 12.2, that the isostematic wavelength λ_s is different at different parts of the sea. Conversely, therefore, it may be said that isostematism with different radio frequencies occurs at different parts of the sea's surface. The interfering signals on the two radio transmissions do not therefore originate at the same place, and no exact relationship between their beat records may in consequence be expected.

(4.2) The Sea as a Random Surface

The problem may be viewed another way, based on the general concept of the sea as a random surface.[3] The sine waves which are the components of its complex configuration lie in a wide and continuous spectrum, and are random in amplitude and phase, as well as in direction. Movement of the waves in one general direction may be considered as the movement of a group of unrelated sine waves, each having its own characteristic velocity. The faster waves travel ahead while the slower fall back, and the group as a whole spreads.[4] If the profile of the group in the plane normal to the wavefront and the sea's surface is examined, it will be seen to be changing continuously, no waveform retaining its identity for very long.

A single-frequency electromagnetic wave falling on such a surface is returned to the receiving aerial as a phase-coherent reflection from the area where the wavelength is predominantly isostematic. A short time later, the dominant wavelength in the area is isostematic with a different frequency, and this frequency is then returned with the greater amplitude. Thus reflection of two frequencies may occur at the same place, but after an interval of time.

There is evidence in the recordings of the interference on the sound and vision frequencies of a smooth transition of the effect from one transmitted frequency to the other. In certain sections of the chart an alternation in the amplitude of the beats at the two frequencies may be clearly seen, the duration of each being about 20 sec. An example is reproduced in Fig. 3(c). This may be explained easily by the spreading of a group of sea waves already described.

The reflecting area for the interference at Kingsand which, in the geometrical idealization of Fig. 9, is shown at C for the vision frequency, must therefore be visualized as oscillating randomly along the dashed line through C. The reflecting area for the sound frequency would oscillate in a like manner, but the relative positions of the two areas would vary randomly.

(4.3) Transient Response of an Isostematic System

Reference has already been made to the fact that the interference, when observed on the waveform monitor, was greatest at the commencement of a line. The sudden rise in carrier amplitude at the end of the line-synchronizing pulse implies that a wide spectrum of frequencies is being radiated at that instant. The distortion of the trace on the monitor, such, for instance, as that depicted in Fig. 5(b), is analogous to the distortion which occurs when the same pulse is passed through a narrow-passband filter.

A characteristic of a narrow-band filter is that the phase angles by which two signals passing through it are delayed are not in the ratio of their frequencies. The sea's characteristic as a filter, when its wavelength is isostematic with the wavelength of the radio carrier, may be examined, at least to a first order of approximation, by considering the phase of the back-scattered field at the receiving aerial relative to that of the diffracted field. Let f_s, representing the diffracted field from North Hessary Tor, be the isostematic frequency; and let the other frequency be f_1, later to be associated with the back-scattered field. At frequency f_1 the phase difference between the parcels of energy scattered back at two consecutive wavecrests is $\pm 2\pi f_1/f_s$, since it is 2π at the isostematic frequency. Over an interval of n wavecrests it is $\pm 2\pi n f_1/f_s$. If this phase difference is compared with a signal at the isostematic frequency, the relative phase is given by the expression

$$\text{Phase difference over } n \text{ wavecrests} = 2\pi n \pm 2\pi n \frac{f_1}{f_s} \quad (2)$$

Consider now what takes place when the large band of frequencies generated at the end of the line-synchronizing pulse falls on a limited area of the sea's surface extending over a few wavecrests only. Radiation at each of the frequencies incident on the area and deflected to the receiving aerial may be represented by an elementary vector whose phase at the receiving aerial depends on the length of the path traversed. The path length is in such a case a step-function, the magnitude of the steps depending on the spacing between wavecrests. The phase of the resultant of the vectors therefore depends on the position of the back-scattering area and on the wavelength of the sea waves.

Elementary vectors representing the energy back-scattered at other small areas are also components of the received signal. They have phase distributions and mean path lengths which are in each case different, so that the phases of their resultants also are different. It is clear, therefore that the phase of the back-scattered energy at the receiving aerial is a continuous function of position.

By definition, the phase of an isostematically back-scattered field is identical with that of a direct (or diffracted) field at the receiving aerial. If the phase difference between the vectors representing the back-scattered field and the diffracted field from the transmitter is given the value $2\pi n(1 \pm f_1/f_s)$, f_1 may be regarded as the 'effective' back-scattered frequency.

Returning again to Fig. 5(b), it is clear that at the peak of the half-cycle above AB, the two frequencies f_s and f_1 are in phase at the receiving aerial. At the instants A and B they are in quadrature because these points lie on the axis of the ripple. Let f_1 lead f_s by $\frac{1}{2}\pi$ radians at A; then at B it lags by an equal amount, and the whole difference in phase across the line AB is π radians. This line has a length equivalent to about 6 microsec of delay.

The back-scattered energy received at instant B has taken a path 6 microsec longer in time, or 1 800 m longer in distance, than the energy received at instant A. Substantially the whole of this difference in path length lies in the distance between the scattering areas and the receiving aerial. If it extends over n wavecrests,

$$\pi = 2\pi n \left(1 \pm \frac{f_1}{f_s}\right)$$

which gives

$$f_1 = f_s \left(1 \pm \frac{1}{2n}\right)$$

Using

$$\lambda_s = 0 \cdot 6\lambda = 3 \cdot 5 \, \text{m}$$

$$n = \frac{1\,800}{3 \cdot 5} = 515$$

and

$$f_1 = 51 \cdot 75 \pm \frac{51 \cdot 75}{1\,030} \, \text{Mc/s}$$

$$= 51 \cdot 75 \, \text{Mc/s} \pm 50 \, \text{kc/s}.$$

This is the bandwidth at 6 dB points, and represents quite a narrow-band filter. The width of the filter clearly depends only on the extent of the sea surface over which isostematism persists.

(4.4) Comparison with a Diffraction Grating

It is clear that isostematism at radio frequencies resembles closely the optical properties of a diffraction grating. As with grating phenomena, the requirement essential to the mechanism here considered is periodicity in the configuration of the irradiated surface. A region of the sea's surface whose ridges are regularly spaced and appropriately oriented with respect to the geometry and frequency of the transmitter/receiver circuit

behaves like a diffraction grating moving towards
or away from the receiver. In doing so, it produces
interference beats on the original transmission which
result from a Doppler change in the frequency of
the energy returned from the sea. In the case of the
sea, no less than in the case of the grating, the ampli-
tude and bandwidth of the reflected field are dependent
only on the total number of lines contained in the
grating.

(5) DEPENDENCE OF BACK SCATTERING ON AERIAL HEIGHTS

It is of interest to examine the way in which
changes in the heights of the transmitting and re-
ceiving aerials affect the amplitude of the back-scattered
signal. Reports from Kingsand were consistent in
saying that the interference was less severe when, in
the period of the temporary transmission, the radiating
aerial was only 150 ft above ground level.

The manner in which the problem is treated here
has been suggested by a paper by Wiltse, Schlesinger
and Johnson,[5] who, following radar practice, have
used the concept of the back-scattering cross-section
of the sea.

A well-known radar equation may be written

$$P_r = k\frac{P_t}{r^2}\frac{\sigma}{r^2} \quad . \quad . \quad . \quad . \quad (3)$$

where P_t and P_r = Transmitted and received powers.
   r = Distance of the target from the radar station.
   σ = Back-scattering cross-section of the target.
   k = A constant.

In the case of a broadcast transmission the transmitter and
receiver are not on a common site. Specifically, in the present
example, the transmitter is remote from the sea. Though, as
will be seen later, the distance between it and the scattering area
varies with changes in the height of the transmitting aerial, the
differences are small compared with the total distance. No
great error is introduced by treating the factor P_t/r_2 in eqn. (3)
as constant. The equation therefore becomes

$$P_r = k_1\frac{\sigma}{r^2} \quad . \quad . \quad . \quad . \quad . \quad (4)$$

where r is the distance from the scattering area to the receiving
aerial and k_1 is a new constant.

The quantity σ is dependent on the physical area of the back-
scattering surface, the electrical properties of sea water and the
statistical distribution of the elementary scattering surfaces. If
these are expressed by A, ρ and s, respectively, and if k_2 is a
constant,

$$\sigma = k_2 A\rho s \quad . \quad . \quad . \quad . \quad . \quad (5)$$

It is here assumed that the angle of the receiving aerial beam is
wide enough to receive always the whole of the back-scattered
energy.

The lengths of the paths from the scattering area to the trans-
mitting aerial at North Hessary Tor and the receiving aerial at
some point on the coast are very much greater than the aerial
heights. Changes in the angle subtended at the sea by the
transmitting and receiving aerials when their heights are varied
are therefore very small, of the order of a few minutes of arc.
It may in consequence be assumed that in the present problem
no change of statistical distribution or of the electrical behaviour
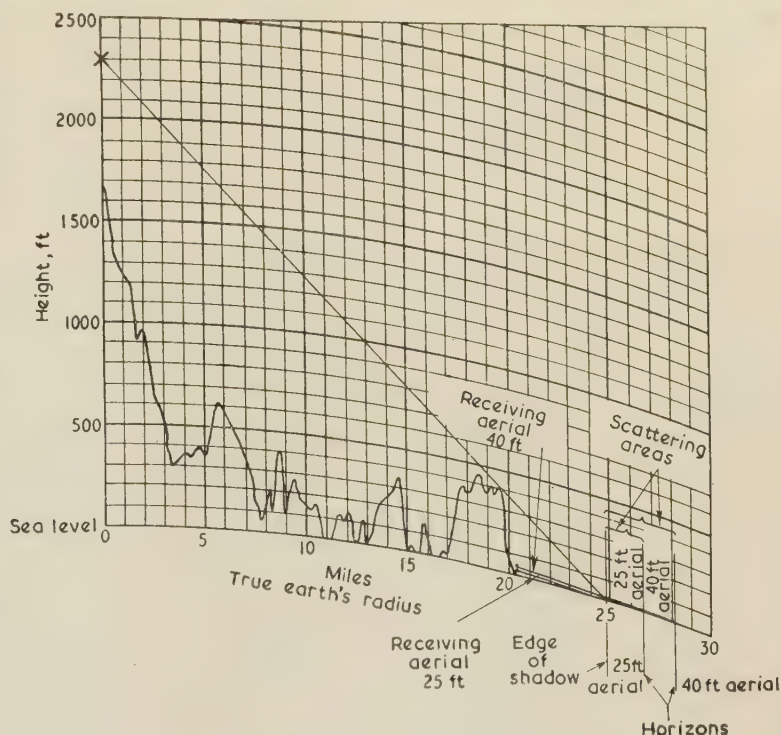

Fig. 10.—Profile of the propagation path between North Hessary Tor
and Downderry.

of the sea occurs over the whole possible range of heights. The
quantities ρ and s in eqn. (5) thus become constants, and the
equation may be restated:

$$\sigma = k_3 A \quad . \quad . \quad . \quad . \quad . \quad (6)$$

The back-scattered power is then

$$P_r = k_1\frac{k_3 A}{r^2}$$

$$= k_4\frac{A}{r^2} \quad . \quad . \quad . \quad . \quad . \quad (7)$$

i.e. it is proportional to the scattering area and inversely pro-
portional to the square of its distance from the receiving aerial.

Fig. 10 illustrates the profile of the terrain between North
Hessary Tor and the site at Downderry used during the tests.
The sea in front of this site is unobstructed over a wide arc
and capable of being represented more closely by a geometrical
idealization than Plymouth Sound. It is for this reason used as
the basis of the following analysis.

The cliff which shades the receiving aerial at Downderry
from the direct transmission is 420 ft above sea level and 19·6
miles from North Hessary Tor. The edge of the shadow made
by this cliff is the near-side limit of the scattering area of the
sea. The far-side limit is the horizon of the receiving aerial.
The area between is A, which, in the calculation, has been
assumed to be proportional to the distance between the limits; i.e.

$$A = k_5(r_h - r_e) \quad . \quad . \quad . \quad . \quad (8)$$

where r_h and r_e are the distances from the receiving aerial of
the horizon and the shadow's edge respectively, and k_5 is a
constant.

The distance of the area A from the receiving aerial is suffi-
ciently well represented by the r.m.s. distance between r_h and r_e.

Calling this r_0, the equation for the back-scattered power may now be written

$$P_r = k_4 \frac{k_5(r_h - r_e)}{r_0^2}$$

$$= k_6 \frac{r_h - r_e}{r_0^2} \quad . \quad . \quad . \quad . \quad . \quad (9)$$

In the present case, r_0 varies over a range 5·3–6 miles for the lower of the two receiving heights considered and 6·1–7·5 miles for the higher. These variations are small compared with the distance of the scattering area from the transmitting aerial. The power incident on the scattering area therefore remains substantially constant for all transmitting aerial heights, as already assumed.

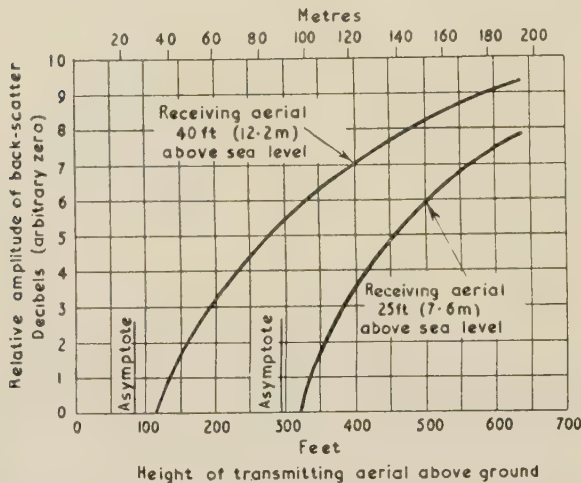The curves of Fig. 11 are drawn for receiving aerial heights

Fig. 11.—Theoretical estimate of the variation in the amplitude of the back-scattered signal with changes in the transmitting and receiving aerial heights.

which are 25 ft and 40 ft above sea level. They express in decibels, with reference to an arbitrary zero, the quantity $(r_h - r_e)/r_0^2$ as a function of transmitting aerial height. Each is asymptotic to a line parallel to the ordinate. The transmitting aerial height at which the asymptote cuts the abscissa is that at which the shadow's edge and the horizon of the receiving aerial are coincident. At transmitting aerial heights which lie to the left of the asymptotes, the irradiated area is out of the range of the receiving aerial, and no back-scattering would be observed. It is clear, therefore, that the amplitude of the observed back-scatter may be expected to increase when either aerial is raised.

The curve in Fig. 11 relating to a receiving aerial height of 40 ft indicates that the expected increase in the amplitude of the interfering signal is over 7 dB when the transmitting aerial height is changed from 150 ft to 633 ft above ground level. Treating the direct transmission to the receiving aerial as taking a path diffracted over a straight edge, the calculated reduction in the shadow loss for the same change in transmitter aerial height is only 1 dB. It is therefore probable that the interference/signal ratio increased by about 6 dB at Downderry when the low temporary aerial system at North Hessary Tor was superseded by the high system at present in use.

(6) CORRELATION WITH METEOROLOGICAL DATA

Many factors are active in producing waves in the sea, and each contributes a characteristic band of wavelengths,

The calculations made in Section 4 show that the sea waves which bring about the interference at Kingsand have wavelengths of about 3·5 m. Waves of this order of magnitude are generated by meteorological factors, and an examination of meteorological data indicates that two factors are closely associated with the interference experienced at Kingsand, one more so than the other. They are wind direction and changes of barometric pressure, respectively.

(6.1) Effect of Wind

References[4, 6] are given in Section 11 in which the relationships between wind and waves are discussed.

The effect of wind on the sea is more directly illustrated in Figs. 12 and 13A, which are aerial photographs of Plymouth Sound. The speed and direction of the wind prevalent when the photographs were taken are indicated. The crests of waves generated by the wind are normal to its direction, and it is easy to pick out in each case a system of waves so oriented. As an indication of scale, a line representing 100 m is drawn in each Figure. This length would span 29 waves, if the value $\lambda_s = 3 \cdot 5$ m derived in Section 4 is assumed to apply to this part of the sea. To examine whether a pattern of such fine structure is present in the photographs, an enlargement of part of Fig. 13A is shown in Fig. 13B. The ratio of scales is about four. The 100 m line appears again in the enlargement, and some 27 waves may be counted along its length. It is thus seen that wind-generated waves are of the right order of magnitude to produce isostematic reflection.

(6.2) Observations on Reception compared with Meteorological Data

Isostematic reflection occurs when the wavecrests, suitably spaced, lie tangentially to ellipses based on the transmitter and receiver as foci. It is seen from Fig. 9 that this condition is possible at Kingsand when the wind direction falls within the ranges 105°–160° and 285°–340°, but the boundaries are not sharply defined. There is a wide spread in the direction of sea waves, and wind directions a little outside the stated arcs may be taken as capable of producing waves of the right orientation.

Local observers at Kingsand had noted that three consecutive days, 22nd–24th June, were unusually free from interference. This period is therefore chosen for examination. Wind data and barometric tendencies for the period 21st–27th June from the Daily Weather Report of the Meteorological Office are reproduced in Table 1. Barometric tendency, which describes the pressure characteristics for the six-hour period prior to the time of the entry, is recorded in code in the weather reports. A key is published in the introduction to the reports issued at the beginning of each quarter. For the present comparison, barometric tendency has been grouped into two classes, 'steady' and 'unsteady', signified in Table 1 by the symbols S and U. These classes may be broadly defined as follows:

When no change, or a change in one direction only has occurred in the previous six hours, the tendency is classed as steady. When, during the six-hour period, the barometer has both risen and fallen, the tendency is classed as unsteady.

Below Table 1 are notes on reception conditions for the days in question. They are reproduced from entries made in his diary by the Engineer-in-Charge of the B.B.C.'s station at Plymouth.

In Table 1, the wind directions shown in bold-face type are those which lie within the ranges discussed at the beginning of this sub-section. Directions lying just outside these ranges are shown in italics.

From these data it may be concluded that interference would be present all the time on the 21st and 26th June, and this is
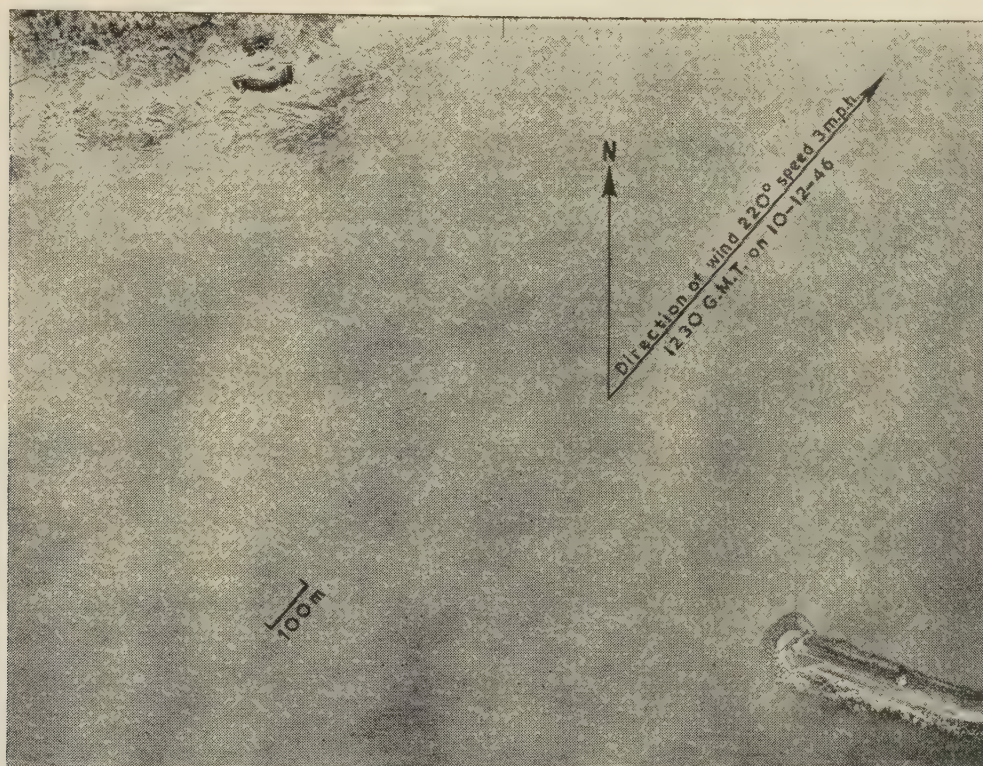
Fig. 12.—Aerial view of Plymouth Sound, taken around noon on the 10th December, 1946.
Measurements of wind speed and direction were not made at Plymouth on this date, and the values quoted are of measurements made at Lizard.

Table 1

WIND DATA AND BAROMETRIC TENDENCIES, 21ST–27TH JUNE

Date June, 1956	0100 hours B.S.T.			0700 hours B.S.T.			1300 hours B.S.T.			1900 hours B.S.T.		
	Wind Direction	Speed knots	Barometer	Wind Direction	Speed knots	Barometer	Wind Direction	Speed knots	Barometer	Wind Direction	Speed knots	Barometer
21st	**337°**	7	U	**326°**	7	S	**337°**	12	S	*348°*	5	U
22nd	45°	6	S	202°	2	U	236°	11	U	0°	0	S
23rd	0°	0	S	0°	0	U	225°	10	S	225°	4	S
24th	0°	0	U	101°	2	S	202°	6	S	0°	0	S
25th	0°	0	U	0°	0	S	*169°*	6	U	220°	2	S
26th	*281°*	4	S	**310°**	6	U	220°	10	U	**326°**	12	U

Notes on Reception Conditions

June 21st.—Interference present at average strength.
June 22nd.—No interference on evening transmission. (No observations were made earlier in the day.)
June 23rd.—Observations from 1200 hours: no interference all day.
June 24th.—Observations at 1500 hours and 1800 hours: no interference. Average interference returns at 2100 hours.
June 25th.—Average interference at 1100 hours; less at 1500 hours; clear of interference by 1700 hours.
June 26th.—Strong interference at 1100 hours, reduced at 1230 hours. At 1430 hours and again at 1700 hours, moderate interference.

borne out by the notes. Again from the data for the 25th June, interference might be expected at 1300 hours, and good reception at 1900 hours. The observations for that day fulfil this expectation exactly.

Short-term variations in atmospheric pressure also appear to induce conditions favourable for interference. For instance, at 1300 hours on the 26th June, interference was present with unsteady pressure, although wind direction was unfavourable. The improvement recorded at 1230 hours on that day might be attributed to the change in the direction of the wind from 310° to 220°, and the change back to 326° at 1900 hours would account for the increase to 'moderate' in the afternoon.

Though the evidence in Table 1 and the notes is not sufficient

to indicate unequivocally the effect of wind velocity, the data for the 26th June are of interest in this respect. The degree of isostematism present at any time is dependent on how nearly the pattern of the sea waves approaches perfect regularity. Greater regularity may be expected with low wind velocities than with high, and it may be concluded that interference will be more severe when the wind velocity is low. The entries in the notes for the 26th June lend some support to this. Interference was strong in the morning of that day and moderate in the afternoon. Though there was little difference in the wind directions at these times, the velocity in the afternoon was twice as great as in the morning.
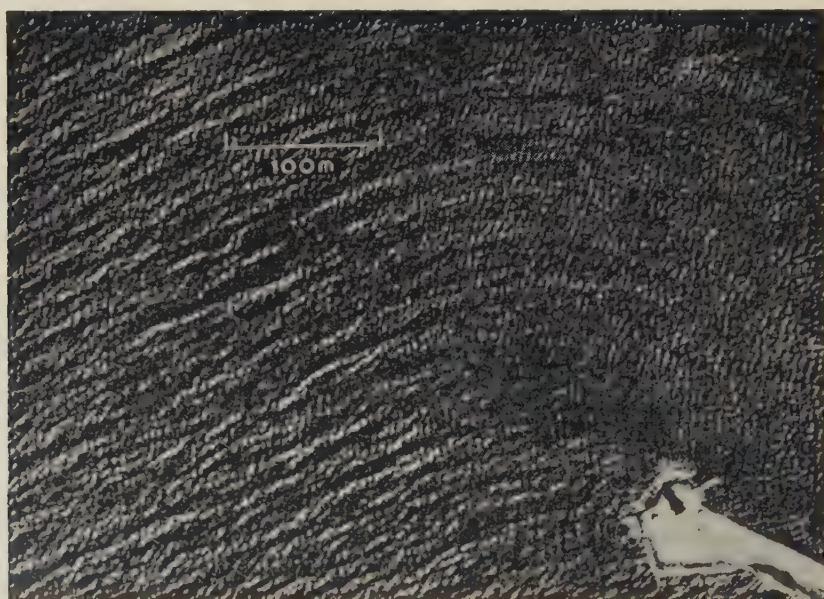
The data in Table 1 accord well with the observations at times

Fig. 13A.—Aerial view of Plymouth Sound, taken around noon on the 28th May, 1947.

Measurements of wind speed and direction were not made at Plymouth on this date, and the values quoted are of measurements made at Lizard.

Fig. 13B.—Part of Fig. 13A enlarged.

Table 2

Date July, 1956	0100 hours B.S.T.			0700 hours B.S.T.			1300 hours B.S.T.			1900 hours B.S.T.		
	Wind		Barometer	Wind		Barometer	Wind		Barometer	Wind		Barometer
	Direction	Speed knots		Direction	Speed knots		Direction	Speed knots		Direction	Speed knots	
17th	**292°**	8	S	0°	0	S	180°	12	S	**146°**	6	S
18th	**124°**	11	U	90°	6	U	**123°**	7	S	*101°*	7	S
19th	**303°**	6	U	**315°**	1	S	191°	2	U	*348°*	7	S
20th	45°	6	S	**337°**	10	U					5	S

of good reception. At 1900 hours on the four days 22nd–25th June, wind direction was unfavourable for interference, and atmospheric pressure was steady. On all these days reception at Kingsand was free from interference.

(6.3) Further Comparisons

Meteorological data for the period occupied by the tests reported here are reproduced in Table 2. They indicate that no interference might be expected at 1300 hours on the 19th July, but that interference would be present at nearly all other times. In fact, interference varied from average to strong during observations throughout the period, except in the afternoon and early evening of the 19th July, when it was slight. By late

tion in planning a service for an area which included coastal towns, especially if one polarization was found to be predominantly better than the other in this respect. An experiment to determine the coefficients would therefore be useful.

To obtain a more detailed picture of the factors which con-

Table 3

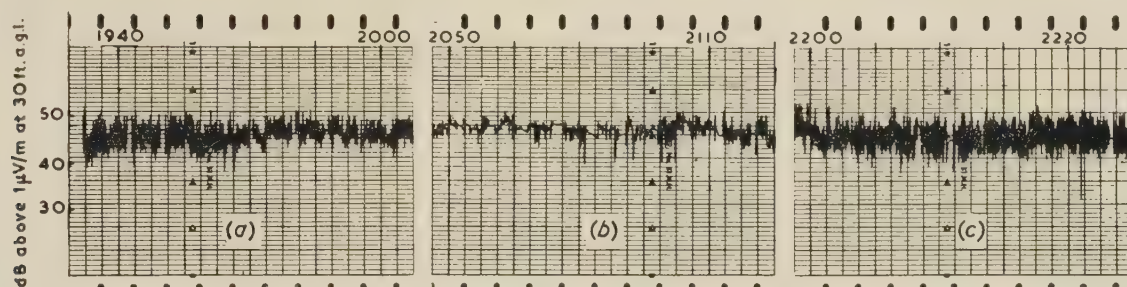| | | Wind | | Barometer |
		Direction	Speed knots	
June 10th	1900 hours B.S.T.	360°	10	S
	2200 hours B.S.T.	340°	10	*
June 11th	0100 hours B.S.T.	326°	2	S

* No reading available.

Fig. 14.—Record demonstrating the sunset effect, taken at Kingsand 10th June, 1956

(a) 1½ hours before sunset.
(b) Immediately before sunset (sunset at 2126 hours B.S.T.).
(c) 1 hour after sunset.

evening on that day interference approached the average. Though from Table 2 this might be expected to have occurred by 1900 hours, interference was not significant until an hour and a half after that time. Nevertheless, correlation is good.

(7) SUNSET EFFECT

An effect which had been noted by local observers and which occurred during the preliminary tests made with field-strength-measuring equipment is that, for a period of about half an hour before sunset, the amplitude of the interference abates significantly. A record of this effect, taken at Kingsand, is reproduced in Fig. 14. Its occurrence is not reliable; it did not, for instance, occur during the four days occupied by the tests with the waveform monitor. From the correlations already made, it may be assumed to be due to a short-term change in the wind before sunset. No evidence of such a change is to be found in the available meteorological data for the time of the recording. This is not surprising, however, as the interval between readings is three hours. The data are reproduced in Table 3.

(8) SUGGESTIONS FOR FURTHER WORK

The observations recorded in Section 3.7 are inadequate for providing information whether the coefficient of back-scattering different for waves polarized vertically and horizontally. Such information would be of value to a broadcasting organiza-

tribute to the occurrence of the phenomenon, a programme of work for correlating isostematic reflection with sea and weather conditions would have to be undertaken. Weather measurements would, however, need to be recorded at more frequent intervals than is the case at present.
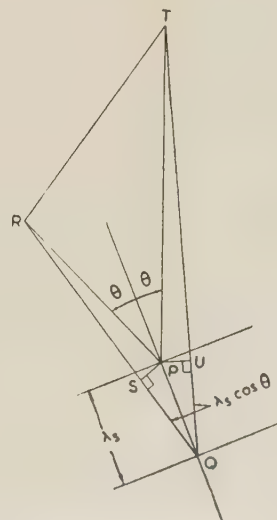

Fig. 15.—Relationship between λ_s and λ.

(9) CONCLUSION

A necessary requirement for the occurrence of the effect discussed here is that the sea within the horizon of the receiving aerial must be directly irradiated by the transmission. This condition is more likely to be met when the transmitter is installed on a site high above sea level and not too distant from the sea. If the receiving aerial has an open aspect towards the irradiated part of the sea and is shielded from the direct transmission by a cliff, the direct and back-scattered signals may have comparable amplitudes and the interference will then be severe.

It is of interest to mention here that the effect is not confined to the transmissions from North Hessary Tor. Since the time of the original tests, isostematic interference has been observed on the Rowridge transmissions at Ventnor, where conditions are similar to those described.

To cause interference, the waves of the sea travelling over the irradiated area must be correctly spaced and suitably oriented with respect to the frequency and geometry of the transmitter/receiver circuit. The energy scattered back then reaches the receiving aerial as a phase-coherent signal of cyclically changing phase, which beats with the direct transmission.

The word 'isostematism' is suggested to express the essential relationship between the phase-coherence of the back-scattered signal and the spatial arrangement of the sea waves. The effect produced may be described as isostematic interference.

There is no easy solution to the problem of reception of television broadcasts in places subject to isostematic interference. In general, the sea is visible over a wide arc in such places, and a directional aerial designed to exclude the whole of the interference would need to have a wide-angle minimum and a front/back ratio of many decibels. Such an aerial, even if easily available, would not infallibly provide a solution in situations where the scattering area of the sea was not in the plane through the transmitting and receiving aerials. With the aerial minimum turned towards the sea, its maximum may in such cases be directed towards a cliff, where strong permanent echoes more objectionable than the interference may originate. One practical way of dealing with the problem would be to mount a communal aerial in a region of high field strength at the top of the cliff, and to distribute the signal to the houses below by cable. The method presents no grave technical difficulties, and the restrictions on its adoption are only economic.

(10) ACKNOWLEDGMENTS

(11) REFERENCES

(1) CROMBIE, D. D.: 'Doppler Spectrum of Sea Echo at 13·56 Mc/s', *Nature*, 16th April, 1955, **175**, p. 681.
(2) DAVIES, T. V.: 'The Theory of Symmetrical Gravity Waves of Finite Amplitude I', *Proceedings of the Royal Society*, 1951, **208**, p. 475.
(3) ECKART, CARL: 'Scattering of Sound from the Sea Surface', *Journal of the Acoustical Society of America*, 1953, **25**, p. 566.
(4) BRETSCHNEIDER, C. L.: 'The Generation and Decay of Wind Waves in Deep Water', *Transactions of the American Geophysical Union*, 1952, **33**, p. 381.
(5) WILTSE, J. C., SCHLESINGER, S. P., and JOHNSON, C. M.: 'Back-Scattering Characteristics of the Sea in the Region from 10 to 50 kMc', *Proceedings of the Institute of Radio Engineers*, 1957, **45**, p. 220.
(6) BURLING, R. W.: 'Surface Waves on Enclosed Bodies of Water', *Proceedings of the Fifth Conference on Coastal Engineering*, September, 1954, Chapter I.

(12) APPENDICES

(12.1) Details of Test Sites

Site	Distance from sea yd	Distance from sea m	Height above sea level ft	Height above sea level m	Height of aerial above ground ft	Height of aerial above ground m	Estimated ambient field strength
Downderry ..	70	64	25	8	25	8	250 μV/m
Kingsand ..	30	27	40	12	25	8	300 μV/m
Staddon Heights	440	400	350	107	25	8	3 mV/m
Plymouth Sound	—	—	—	—	15	5	Not estimated

The field strengths noted here are estimated from the deflection of the a.g.c. meter on the receiver and are very approximate.

(12.2) Relationship between λ_s and λ

TR is the path of the direct ray; TPR and TQR the paths of reflected rays such that

$$TQR = TPR + \lambda$$

$$PQ = \text{Distance between two wavecrests} = \lambda_s$$

$$SQ = UQ = \lambda_s \cos\theta \simeq \tfrac{1}{2}\lambda$$

Hence

$$\lambda_s = \tfrac{1}{2}\lambda \sec\theta$$

and the loci of constant λ_s are circles passing through T and R.

For the points A, B and C in Fig. 9, the calculated beat frequencies are as follows:

Point	θ	$\sec\theta$	$\tfrac{1}{2}\lambda$ m	λ_s	v m/s	c/s
A	27°	1·12	2·9	3·24	2·25	0·69
B	24°	1·09	2·9	3·16	2·22	0·70
C	44·5°	1·4	2·9	4·05	2·51	
After correction (see Section 4.1)					2·76	0·68

DISCUSSION ON

'TRANSVERSE FILM BOLOMETERS FOR THE MEASUREMENT OF POWER IN RECTANGULAR WAVEGUIDES'*

Mr. L. J. T. Hinton (*communicated*): I agree with the author that there are many advantages in using a transverse resistive-film bolometer for the accurate measurement of power at the milliwatt level in rectangular waveguides. This is particularly so at frequencies above 20 Gc/s when the capacitance of the thermistor bead results in significant error.

In Section 5 the author suggests that further development of the film bolometer at higher frequencies would seem well worth while. I feel I should therefore draw attention to a contribution by Dr. J. Collard in 1955.† This described an instrument for the measurement of power in the range 1–100 mW at frequencies in the region of 35 Gc/s, which absorbed all the incident power in a film of uniform surface resistivity.

Direct comparisons between models of the instrument, which is known as an enthrakometer, and milliwatt water calorimeters at powers in the range 10–20 mW have shown that the accuracy of the enthrakometer at 35 Gc/s is about ±5%.

Mr. Lane points out in Section 2 that the film bolometer he has used differs from the enthrakometer in that the film is narrow in the transverse plane. I agree that this arrangement gives greater sensitivity and more nearly ideal substitution of d.c. for microwave power. However, by the use of an impedance bridge containing the resistive film measuring the power and an identical film for ambient temperature compensation, followed by a high-gain l.f. amplifier, very small resistance changes are detectable with the enthrakometer. The measurement of power at a level of 1 mW at 35 Gc/s presents no difficulty. With regard

to a possible error due to a difference in temperature gradient between the r.f. power measurement and the d.c. calibration, I can only state that none has been detected.

The thin-strip film bolometer, on the other hand, does set up higher order modes which might result in the dissipation of power elsewhere than in the resistive film, or in the loss of power through the r.f. decoupling circuits. Also, of course, its intrinsic bandwidth will be less than that of the enthrakometer, due to the presence of the inductive component in its equivalent circuit.

Mr. J. A. Lane (*in reply*): Details of the instrument mentioned by Dr. Collard in his communication have not been published, and I am indebted to Mr. Hinton for providing these results on this type of bolometer. This information confirms the merits of resistive-film techniques for power measurement in rectangular waveguides.

I agree that there is a possibility that higher-order modes may introduce errors with the narrow-film bolometer, and that some reduction in bandwidth results from this arrangement by comparison with the bandwidth obtained when using a bolometer which fills the transverse cross-section. However, each technique has its advantages, as Mr. Hinton points out, and the instrumental error in the method I described appears to be small. In addition, the use of a narrow film facilitates the construction of a very simple indicating circuit. If necessary, the bandwidth can be improved by the addition of a movable plunger following the resistive film.

Subsequent experiments have shown that the narrow-film technique can also be used for powers of the order of a few milliwatts or more at wavelengths as long as 10 cm.

* LANE, J. A.: Paper No. 2488 R, January, 1958 (see 105 B, p. 77).
† Discussion on WILLSHAW, W. E., LAMONT, H. R., and HICKIN, E. M.: 'Experimental Equipment and Techniques for a Study of Millimetre-Wave Propagation' (see 102 B, p. 312).

DISCUSSION ON

'DESIGN OF MICROWAVE FILTERS WITH QUARTER-WAVE COUPLINGS'*

Mr. L. K. Knoch (*communicated*): I was interested in the high values of susceptances.[H, J, K, L] The three post susceptances were used in the 0·9 in × 0·4 in waveguide in the 9·3 Gc/s band. The measurements were made with the post-diameter to waveguide-width ratio d/a lying in the range from 0·019 to 0·069. For two different $2a/\lambda_g$ ratios, graphs were obtained as in Fig. H. The 'left curving' of the d/a plots in terms of normalized susceptance $-b$ or $-b(2a/\lambda_g)$ for the higher values of d/a was not observed.

Up to the value of $d/a = 0·045$ the experimental plots run parallel to the theoretically computed ones at a constant distance. The shift of the experimental plot position against the theoretical one is an obvious function of the $2a/\lambda_g$ ratio. Beyond the value of $d/a = 0·045$ the discrepancy between the experimental and theoretical curves grows larger, certainly for the reasons given by Messrs. Craven and Lewin in their reply.[K] But the experimental plots remained linear up to the highest susceptance values measured.

It is evident from the graphs that the $d/a = f[b(2a/\lambda_g)]$ func-
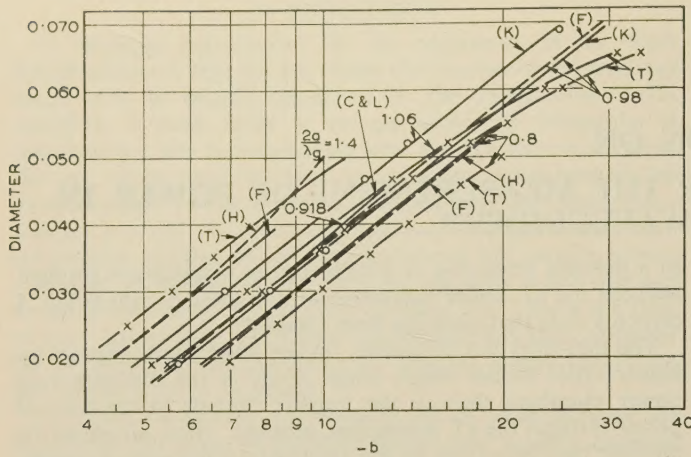
* See Reference H on page 396.

Fig. H

(C & L) Craven and Lewin's measurements.
(H) Horna's measurements.
(K) Knoch's measurements.
(F) Knoch's formula.
(T) Theoretical values.

Fig. J

(C & L) Craven and Lewin's measurements.
(H) Horna's measurements.
(K) Knoch's measurements.

tion follows the relation $d/a = K(2a/\lambda_g) + M \log b$, where $M = 0.71$ (constant) and K is assumed to be a function of $2a/\lambda_g$.

The value of M was taken as a mean value from the measurements made by Messrs. Craven and Lewin, Mr. Horna and myself.

The figures obtained from the above-mentioned measurements permit a simple relation to be established (Fig. J):

$$K = \left[2\left(\frac{2a}{\lambda_g}\right) - 5.4 \right] 10^{-2}$$

One plot for $2a/\lambda_g = 1.06$ failed to follow this relation, and because of a limited number of measurements, it is not possible to conclude whether the formula should hold over the whole range of $2a/\lambda_g$ ratios.

Yet, because of the great usefulness of susceptances of this kind, especially in microwave filter design, it seemed desirable to publish this information.

Messrs. G. Craven and **L. Lewin** (*in reply*): Some interest appears to exist in the properties of triple-post susceptances, and Mr. Knoch's communication and the considerable amount of data he presents provides further evidence for the need for a second-order (in the post diameter) formula for the admittance. By analogy with the second-order theory for the single post we may expect, in addition to a contribution to the post susceptance, a series or transformer effect, though for the triple-post this
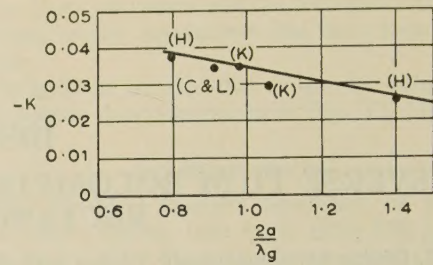
effect is likely to give a minor contribution to the impedance in a considerable range in which the correction terms to the susceptance may be appreciable. The form taken by the formula will not, however, be the same as Knoch's empirical equation. From a comparison with the single-post formula[M] it is apparent that the normalized susceptance b will be related to d by an equation of the form

$$\frac{1}{b} = K \log \frac{d}{a} + \frac{d^2}{a^2} \text{f}$$

where f is a function of a, λ and $\log (d/a)$.

The representation of d/a as a linear function of $\log b$ in a certain range of the variables, as demonstrated by Knoch, is therefore a convenient numerical coincidence, as may be met also in the theory of coaxial exponential transmission lines, but it has no special analytic significance.

One of the difficulties likely to be met in working out a second-order theory is the need to take adequate account of the proximity effect between the posts, and this difficulty becomes accentuated, even for small post diameters, as the number of posts under consideration increases.

REFERENCES

(H) CRAVEN, G., and LEWIN, L.: *Proceedings I.E.E.*, Paper No. 2001 R, March, 1956 (**103 B**, p. 173).

(J) LEVY, R.: Discussion on Reference H, *ibid.*, 1956, **103 B**, p. 632.

(K) HORNA, J.: Discussion on Reference H, *ibid.*, 1957, **104 B**, p. 528.

(L) SIMON, J. C., and BROUSSARD, G.: 'Les filtres passe-bands en hyperfréquence', *Annales de Radioélectricité*, 1953, **8**, p. 31.

(M) LEWIN, L.: 'Advanced Theory of Waveguides' (Iliffe and Sons, 1951), p. 35.

PROCEEDINGS OF THE INSTITUTION OF ELECTRICAL ENGINEERS

CONTENTS

THE BENEVOLENT FUND

The number of applications for assistance from the Fund has shown a marked increase during the last few years, and this year these fresh demands exceed the increase in contributions. The state of the Fund has enabled the Court of Governors to maintain for the present their standard of assistance in the necessitous cases but they are anxious that their ability to help should not be impaired.

The Fund is supported by about a third of the members, and the Governors' best thanks are accorded to those who subscribe. They do, however, specially appeal to those who do not at present contribute to the Fund to do so, preferably under deed of Covenant.

Subscriptions and Donations may be sent by post to

THE INCORPORATED BENEVOLENT FUND OF
THE INSTITUTION OF ELECTRICAL ENGINEERS
SAVOY PLACE, LONDON, W.C.2

or may be handed to one of the Local Hon. Treasurers of the Fund.

THE FUND IS SUPPORTED BY SUBSCRIPTIONS, DONATIONS, LEGACIES

LOCAL HON. TREASURERS OF THE FUND:

EAST MIDLAND CENTRE . . . . . R. C. Woods	SCOTTISH CENTRE . . . . R. H. Dean, B.Sc.Tech.
IRISH BRANCH . . . . A. Harkin, M.E.	NORTH SCOTLAND SUB-CENTRE . . . P. Philip
MERSEY AND NORTH WALES CENTRE D. A. Picken	SOUTH MIDLAND CENTRE Capt. J. H. Patterson, R.A.
NORTH-EASTERN CENTRE. . . J. F. Skipsey, B.Sc.	RUGBY SUB-CENTRE . . . . . . P. G. Ross, B.Sc.
NORTH MIDLAND CENTRE E. C. Walton, Ph.D., B.Eng.	SOUTHERN CENTRE . . . . . . G. D. Arden
SHEFFIELD SUB-CENTRE . . . . . F. Seddon	WESTERN CENTRE (BRISTOL) . . A. H. McQueen
NORTH-WESTERN CENTRE . E. G. Taylor, B.Sc.(Eng.)	WESTERN CENTRE (CARDIFF) . . E. W. S. Watt
NORTH LANCASHIRE SUB-CENTRE	WEST WALES (SWANSEA) SUB-CENTRE O. J. Mayo
G. K. Alston, B.Sc.(Eng.)	SOUTH-WESTERN SUB-CENTRE . . W. E. Johnson
NORTHERN IRELAND CENTRE . . G. H. Moir, J.P.	